Weak Lensing Galaxy Cluster Mass Estimates in the Dark Energy Survey

Tamás Norbert Varga

München 2020

Weak Lensing Galaxy Cluster Mass Estimates in the Dark Energy Survey

Tamás Norbert Varga

Dissertation
an der Fakultät für Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Tamás Norbert Varga
aus Budapest

München, den 20. Juli 2020

Erstgutachter: Prof. Dr. Ralf Bender
Zweitgutachter: Prof. Dr. Jochen Weller
Tag der mündlichen Prüfung: 25. September 2020

Contents

List of Figures

List of Tables

Zusammenfassung

Die Untersuchung von Galaxienhaufen eröffnet einen wichtigen Weg zum Verständnis der Zusammensetzung und Entwicklung des Universums und seines verborgenen dunklen Sektors. Da Galaxienhaufen zu einem großen Teil aus dunkler Materie bestehen, geben praktische Methoden zu ihrer Entdeckung in Beobachtungsdaten keine direkte Information über ihre Gesamtmassen, und es wird notwendig, andere Methoden wie die Gravitationslinsentechnik zu verwenden, um auf ihre Massenverteilung zu schließen.

Im ersten Teil der Arbeit erstelle ich eine Kalibration der Beziehung von beobachtbarer Galaxienzahl zur Gesamtmasse für Galaxienhaufen im mittels des schwachen Gravitationslinseneffektes. Die Daten hierfür wurden durch die Dark Energy Survey gesammelt und reduziert. Desweiteren entwickle ich Methoden zur Quantifizierung und Verbesserung der Robustheit und systematischen Unsicherheit von Messungen des schwachen Gravitationslinseneffektes. Diese Kalibrierung der Massen von Galaxienhaufen ist ein entscheidendes Bindeglied, das die kosmologische Theorie mit der beobachteten Realität verbindet. Sie ermöglicht es, Sie ermöglicht es, kosmologische Schlüsse aus den Massen und Häufigkeiten von Galaxienhaufen zu ziehen. Darüber hinaus untersuche ich eine Art von systematischem Fehler bei der Schätzung des schwachen Gravitationslinseneffektes von Galaxienhaufen, die mit der potenziell falschen Schätzung der Linsengeometrie aufgrund der Kontamination des Quellgalaxienkatalogs durch Galaxien, die mit dem Zielhaufen assoziiert sind, zusammenhängt. In diesem Zusammenhang validiere ich eine Korrekturmethode, die sich auf die geschätzte photometrische Rotverschiebungs-Wahrscheinlichkeitsverteilungsfunktion der Quellgalaxien stützt.

Schließlich konstruiere ich synthetische optische Beobachtungen von Galaxienhaufen, mit denen die Messmethoden in einer vollständigen Art und Weise von simulierten Bilddaten bis zu wiedergefundenen Haufenmassen verglichen werden kann. Um dies zu erreichen, erstelle ich ein transparentes statistisches Modell unter Verwendung eines datengesteuerten Ansatzes zur Messung und Beschreibung der photometrischen Eigenschaften von Galaxienhaufen und ihrer Umgebung und erstelle simulierte Beobachtungdaten von Galaxienhaufen, die die reale Verteilung von Galaxien und ihre beobachteten photometrischen Eigenschaften imitieren und extrapolieren.

Abstract

The study of galaxy clusters opens a prominent pathway towards understanding the composition and evolution of the Universe and its hidden dark sector. As galaxy clusters are to a large extent composed of dark matter, practical methods to detect them in observational data do not give direct information about their total masses, and it becomes necessary to make use of other methods such as gravitational lensing to infer their mass distributions.

In the first part of the thesis I perform a weak lensing calibration of the observable - mass relation for galaxy clusters in the optical observational dataset collected and reduced by the Dark Energy Survey, and develop the methods to quantify and improve the robustness and systematic uncertainty of weak lensing measurements. This cluster mass calibration is a crucial link connecting and anchoring cosmological theory to the observed reality, and facilitates the ongoing efforts to derive cosmological constraints from the masses and abundances of galaxy clusters. I furthermore investigate a type of systematic bias in cluster weak lensing estimates which relates to the potentially incorrect estimation of the lens-geometry due to the contamination of the source galaxy catalog by galaxies associated with the targeted cluster. In relation to this I validate a correction method which relies on the estimated photometric redshift probability distribution function of source galaxies.

Finally I construct synthetic optical observations of galaxy clusters, which can be used to benchmark weak lensing measurement in an end-to-end fashion from mock observations to recovered cluster masses. To achieve this I create a transparent statistical learning model using a data driven approach to measure and model the photometric properties of galaxy clusters and their sky-environments, and generate mock galaxy cluster observations which mimic and extrapolate the real distribution of galaxies and their observed photometric properties.

Chapter 1

Introduction

Over the course of the last decades it has been revealed that most of the Universe is composed of previously unexpected forms of matter and energy, aptly named as *dark matter* and *dark energy*, which presents our physical understanding of the cosmos with a new frontier often referred to as the dark sector of the Universe (Bertone & Hooper, 2018; Huterer & Shafer, 2018). While phenomenological descriptions are abound, the nature of these components is as of yet unclear.

The research direction of this thesis is set at investigating the matter distribution and masses of astrophysical objects known as *galaxy clusters*. Clusters are to a large extent composed of dark matter, and through their abundance and growth since the dawn of cosmic history they are informative about the composition and evolution of the Universe and its hidden dark sector (Allen et al., 2011; Weinberg et al., 2013). Hence the study of galaxy clusters has in recent years become a prominent pathway towards understanding the nonlinear growth of cosmic structure (Kravtsov & Borgani, 2012), and in constraining the cosmological parameters of the Universe (Dodelson et al., 2016). While the practical methods to detect galaxy clusters in observational data are different between different observation strategies, available wavelengths and targeted redshifts ranges, they do not give direct information about cluster masses, therefore it becomes necessary to calibrate a so-called *observable - mass relation* (Weinberg et al., 2013).

Weak gravitational lensing on the other hand provides a practical and robust method to study the mass properties of galaxy clusters (Bartelmann & Schneider, 2001). It relies on estimating the shapes or ellipticities of background *source* galaxies, which show an apparent preferential tangential alignment around massive *lens* objects such as galaxy clusters. This preferential alignment is induced by the gravitational potential of the lens system, therefore, lensing is sensitive to the total mass of an object, not just its baryonic component, and is insensitive to the dynamical state of the mass distribution. Furthermore, weak lensing mass measurements are readily scalable to a statistical ensemble of targets in wide field surveys, and for this reason the lensing based calibrations of the observable-mass relations became a standard practice for galaxy cluster based cosmological analyses. However, in spite of the great invested effort, current cosmological constraints are still dominated by uncertainties in the calibration of cluster masses (Rozo et al., 2010; Mantz et al., 2015; Planck Collaboration et al., 2016; Bocquet et al., 2019).

The aim of this thesis is to perform a weak lensing calibration of the observable - mass rela-

tion for galaxy clusters in the optical observational dataset collected by the Dark Energy Survey (DES, DES Collaboration et al., 2016), and to develop the methods to quantify and improve the robustness and systematic uncertainty of the needed weak lensing measurements. This cluster mass calibration measurement is a crucial link connecting and anchoring cosmological theory to the observed reality, and facilitates the ongoing efforts within DES to derive cosmological constraints from the masses and abundances of galaxy clusters (DES Collaboration et al., 2020).

The thesis is divided into three chapters of original research:

1. In the first part of original research presented in chapter 2 and published in McClintock & Varga et al. (2019), we use weak lensing to measure the projected mean galaxy cluster mass profiles of galaxy clusters identified in DES Year 1 (Y1) data via the redMaPPer algorithm (Rykoff et al., 2014) across different samples selected by redshift and richness – an optical mass proxy. We then develop a modeling framework to derive mean cluster masses from the projected profiles, and use these masses to calibrate the mass–richness–redshift relation. These masses and the observed cluster abundances are then used by DES Collaboration et al. (2020) to derive cluster cosmology constraints from the DES Y1 dataset.

2. In the second part of original research, presented in chapter 3 and published in Varga et al. (2019), we investigate a type of systematic bias in cluster weak lensing estimates which relates to the potentially incorrect estimation of the lens-geometry due to the contamination of the source galaxy catalog by galaxies associated with the targeted cluster. We estimate the performance of a correction method for this effect which uses the estimated photometric redshift probability distribution function of source galaxies, and use it to validate cluster mass profile measurements obtained in chapter 2.

3. Finally, in chapter 4 of the thesis we embark on the task to create a full synthetic testbed for cluster weak lensing measurements. Building upon our experience from the previous chapters, we identify that the interplay between different observational and systematic effects can be best characterized in a unified benchmark which tests the measurement in an end-to-end fashion from mock observations to recovered cluster masses. Constructing such mocks which capture the key observational and physical properties of galaxy clusters in DES is non-trivial. In this last chapter of original research therefore we develop a statistical learning model using a data driven approach to measure the photometric properties of galaxy clusters and their sky-environments. In turn this model is then used to generate plausible mock galaxy clusters and model their line-of-sights in a way which captures the real distribution of galaxies and their observed photometric features. This work hence sets the ground for a dedicated cluster image simulation analysis estimating the performance of past and upcoming cluster mass calibration measurements.

In the following our goal is to briefly recapitulate the theoretical concepts serving the basis of the original research presented in the later chapters of the thesis.

1.1 Standard Cosmological Model

On macroscopic scales the physical reality is described within the framework of the theory of general relativity. The current standard cosmological model is also formulated in this language, and we introduce its relevant ingredients below based on the reviews of Mukhanov (2005) and Hogg (1999)

General relativity describes the world as a four dimensional Lorentzian manifold with a metric tensor $g_{\alpha\beta}$. This manifold is commonly referred to as *space-time*, since it unifies the classical concept of space and time into a single entity. The essence of general relativity is that matter and energy induces a 4-curvature in space-time, reflected in the change of the local metric. In turn the change in metric changes the trajectories of motion available for matter and light (Mukhanov, 2005). Mathematically this can be expressed in the Einstein field equation[1]

$$R^{\alpha}_{\ \beta} - \frac{1}{2}\delta^{\alpha}_{\ \beta} R - \delta^{\alpha}_{\ \beta}\Lambda = 8\pi G\, T^{\alpha}_{\ \beta}\,. \tag{1.1}$$

This equation is formulated for 4-tensor fields, where $R^{\alpha}_{\ \beta}$ and R is the so called Ricci curvature tensor and Ricci scalar respectively which are calculated from the metric tensor $g_{\mu\nu}$ and its first and second order derivatives, $\delta^{\alpha}_{\ \beta}$ is the unit matrix, Λ is the so-called *cosmological constant*, while $T^{\alpha}_{\ \beta}$ is the stress-energy tensor of the matter content of the universe. G is the Newtonian gravitational constant. In principle, knowing $T^{\alpha}_{\ \beta}$, which is completely dependent on the microphysics of matter, the field equation can be solved. The above simple form is however deceptive as solving the Einstein field equation is non-trivial, and exact solutions are only known in special cases with symmetries.

1.1.1 An Expanding Universe

A key epistemological assumption in thinking about the universe is the *cosmological principle*, which states that on sufficiently large scales the properties of the universe are the same for all observers (Mukhanov, 2005). In mathematical terms this means that the universe on large scales is assumed to be homogeneous and isotropic. This principle can be experimentally tested, and current observations about the large scale distribution of galaxies (Alam et al., 2017) and the cosmic microwave background (Planck Collaboration et al., 2018) indeed confirm that the cosmological principle holds on length scales larger than hundreds of Mpc.

The cosmological principle enables us to approach describing the Universe in a perturbative way. One can decouple the question into modeling a large-scale homogeneous and isotropic background cosmology, and into the study of small scale physics of perturbations upon the large-scale background.

In this approach hence one needs only to study the background model to understand the large scale evolution and properties of the universe. The symmetries of this large-scale background allow us to find an exact solution metric for the Einstein equation, which is known as the

[1] Assuming $c = 1$ units following Mukhanov (2005)

Friedmann-Lemaitre-Robertson-Walker (FRW) metric (Mukhanov, 2005). For this metric the infinitesimal line-element is expressed in spherical coordinates as

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \mathrm{d}t^2 - a^2(t)\left[\frac{\mathrm{d}r^2}{1-Kr^2} + r^2\mathrm{d}\Omega^2\right], \tag{1.2}$$

where t is a time parameter, while r and Ω refer to the radial and angular coordinates respectively. K is the spatial curvature parameter which differentiates globally closed spherical ($K > 0$), flat euclidean ($K = 0$) and open hyperbolic ($K < 0$) space-times. The line element decouples into a time-independent spatial component within the square brackets and a time-dependent prefactor denoted by the $a(t)$ *scale factor*. The normalization of the scale factor is arbitrary, and it is customary to choose it to be $a(t = \text{today}) = 1$.

The above metric allows for explicit time evolution of the background cosmology, whereby space can expand or contract as the scale factor changes. Such expansion was indeed discovered by Hubble (1929) based on the apparent recession of galaxies. The detection of cosmic expansion gave rise to the standard *big bang* model of cosmology. This model at present can accommodate most observational facts about the Universe, and can be loosely described as the idea that the current state of the Universe is a product of cosmic expansion. During this expansion out of relatively simple initial conditions of a hotter and denser early state form the elements, stars, planets and galaxies we experience today (Mukhanov, 2005). In this picture the rate of cosmic expansion is characterized by the Hubble parameter defined as

$$H(t) = \frac{\dot{a}(t)}{a(t)} \tag{1.3}$$

where the current measurements indicate that $H_0 \equiv H(t = \text{today})$, the value of the expansion rate at present is approximately on the order of 70 km / s / Mpc based on a wide range of measurements (Riess et al., 2016; Planck Collaboration et al., 2018). Recent measurements furthermore indicate that the expansion of the Universe is in fact accelerating with time.

Finding the actual expansion history can be done by re-expressing Equation 1.1 using the FRW metric from Equation 1.2 and the symmetry considerations of the cosmological principle and solving the *Friedmann equations* (Huterer & Shafer, 2018)

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\sum_i \rho_i - \frac{K}{a^2} + \frac{\Lambda}{3}, \tag{1.4}$$

$$\left(\frac{\ddot{a}}{a}\right) = -\frac{4\pi G}{3}\sum_i (\rho_i + 3p_i) + \frac{\Lambda}{3}, \tag{1.5}$$

where the summation goes over the different constituent species of the universe. ρ_i and p_i refers to the energy density and pressure of a species respectively. Together these terms constitute the diagonal of $T^{\alpha}{}_{\beta\, i}$ which are the only nonzero components allowed by the cosmological principle.

In practice, the different matter and energy species are commonly described by a fluid model, where the energy and pressure is related via an effective equation of state parameter w_i (Huterer & Shafer, 2018), such that

$$p_i = w_i\, \rho_i \tag{1.6}$$

An often used reference energy value is the *critical density* which is the energy density of a spatially flat universe $K = 0$ with $\Lambda = 0$

$$\rho_c(t) = \frac{3H(t)^2}{8\pi G} \,. \tag{1.7}$$

Using this critical density, a density parameter of the different constituent species can be defined as $\Omega_i = \rho_i/\rho_c$. For the case of the cosmological constant this is defined as $\Omega_\Lambda = \Lambda/(3H_0^2)$, while for curvature $\Omega_K = K/(3\,a^2)$. During cosmic expansion different matter and energy species become diluted to a different extent, which can be expressed from Equation 1.4 and Equation 1.5 as

$$\Omega_i(t) = \Omega_i \cdot a(t)^{-3(1+w_i)} \,, \tag{1.8}$$

where Ω_i represents the value of the density parameter in the present time.

A useful quantity in observationally describing the cosmic expansion history is the redshift z experienced by a photon, which arises due to cosmic expansion between the time of emission from a distant source and absorption by the observer (Hogg, 1999):

$$z = \frac{\lambda_{\rm obs}}{\lambda_{\rm emitted}} - 1 \,. \tag{1.9}$$

This kind of cosmological redshift arises in addition to possible proper motion of the emitter or the absorber, and can be expressed from the scale factor at different times as

$$z = \frac{a(t_{\rm obs})}{a(t_{\rm emitted})} - 1 \,. \tag{1.10}$$

Redshifts are a thus preferred way of quantifying look-back time into cosmic history due to their direct connection to an observable quantity.

In the expanding background the distance between two points can be calculated by integrating the line element along the shortest path between the points on a space-like hypersurface (Hogg, 1999). This simple distance definition is denoted as the proper distance $D_P(t)$, which is time dependent due to the cosmic expansion or *Hubble flow*. Based on the form of the metric in Equation 1.2 the expansion can be factored out by defining the *comoving distance* as $D_C = D_P(t)/a(t)$.

These distance measures are defined between events at the same time, therefore for observational applications they have to be modified to be applicable for events observed in the past at redshift z and an observer in the present. For this scenario following Hogg (1999) the comoving distance is given as

$$D_C(z) = \int_z^0 \frac{{\rm d}z'}{H(z')} \,. \tag{1.11}$$

Alternatively, the *transverse comoving distance* between two events at redshift z with apparent angular separation $\delta\theta$ is given by $D_M(z) \cdot \delta\theta$ where

$$D_M(z) = \begin{cases} \frac{1}{H_0\sqrt{\Omega_k}} \sinh(H_0\sqrt{\Omega_k} D_C(z) & \text{for } \Omega_K > 0 \\ D_C(z) & \text{for } \Omega_K = 0 \\ \frac{1}{H_0\sqrt{\Omega_k}} \sin(H_0\sqrt{\Omega_k} D_C(z) & \text{for } \Omega_K < 0 \end{cases} \,, \tag{1.12}$$

and where the possible spatial curvature of space must be taken into account. A further distance metric relevant to this work is the *angular diameter distance* which is used when converting apparent angular $\delta\theta$ separations of events into proper physical separation δx at redshift z, such that $D_A(z) \equiv \delta x/\delta\theta$, or

$$D_A(z) = \frac{D_M(z)}{1+z}\,. \tag{1.13}$$

A curious property of this distance definition is that it is not a monotonically increasing function of redshift, and in fact objects at high redshift appear increasingly larger on the sky. However this is counteracted by the faster than classical decrease of surface brightness, as expressed in the so-called *luminosity distance*

$$D_L(z) = (1+z)\,D_M(z)\,. \tag{1.14}$$

1.1.2 The ΛCDM paradigm

In our standard picture the stress-energy of the Universe is made up by the following types of components (Mukhanov, 2005; Weinberg, 2008; Planck Collaboration et al., 2018):

- **Radiation**: Photons and relativistic particles such as neutrinos with effective equation of state $w = 1/3$ (as defined in Equation 1.6). While at very early times the universe was dominated by radiation, due to the steep dilution of energy density as $\rho_\nu \propto a^{-4}(t)$ it quickly became subdominant during cosmic expansion.

- **Baryonic Matter**: Non-relativistic particles described by standard model physics which interact electromagnetically[2] . Hence all visible matter: stars, planets, dust and gas belong in this category. On very large scales matter is modeled as a pressureless *dust*, hence $w = 0$ and $\rho_m \propto a^{-3}(t)$.The density parameter of baryons today is $\Omega_b \approx 0.04$.

- **Dark Matter**: A type of matter which does not interact electromagnetically, and outweighs traditional baryonic matter in the total energy budget by several factors. On cosmological scales the effective equation of state of dark matter is identical to baryonic matter. Dark matter is often modeled as *cold dark matter* (CDM) which means that it has no intrinsic thermal kinetic energy, and it is further often assumed to be collisionless or non-self interacting. The density parameter of dark matter today is $\Omega_{DM} \approx 0.25$, however the more commonly used parameter is the total amount of matter $\Omega_b + \Omega_{DM} = \Omega_m \approx 0.29$.

- **Dark Energy**: A type of homogeneous energy density responsible for the observed accelerating expansion of the Universe. In the simplest model equivalent to the cosmological constant Λ, in which case $w = -1$ where $\rho_\Lambda \propto a(t)$ that is it does not become diluted with cosmic expansion and as dominant component ($\Omega_\Lambda \approx 1$) drives an exponential expansion of the Universe $a(t) \propto e^{H\,t}$. Current measurements indicate that the density parameter of dark energy today is $\Omega_\Lambda \approx 0.71$.

[2]Hence in a cosmological context electrons are also considered to be baryons

The above outlined energy components define the Λ - cold dark matter (ΛCDM) paradigm. The key feature of this model is that at present times $\sim 96\%$ of the energy budget of the universe is made by the dark sector composed of dark matter and dark energy, and is thus not available for direct observation. Since dark matter does not interact with photons, small initial perturbations in the dark matter field are free to undergo gravitational collapse and form structures even when baryonic matter is still coupled with radiation in the early universe. Hence in the ΛCDM picture cosmic structure is initially formed by dark matter, and only at a later time after decoupling from photons can baryonic matter play a role. Furthermore since today the energy budget is already dominated by dark energy resulting in the accelerating expansion of the Universe. Based on the Friedmann equations the spatial curvature K of the metric can also appear as an effective energy density defined by $\Omega_K = K/(3\,a^2)$, however in our present picture the Universe is measured to be very close to spatially flat[3], hence $\Omega_\Lambda \approx 1 - \Omega_m$

1.1.3 Growth of Structure and the Halo Model

A crucial observational fact in cosmology is that small temperature and density perturbations are already visible in the cosmic microwave background (CMB) at $z \approx 1100$, which is the earliest directly observable state of the universe. These density perturbations can be described as a Gaussian random field for which it is customary to define the density contrast as $\delta = \rho/\overline{\rho}–1$. At the time of the CMB $\delta \sim 10^{-5}$. It is our understanding that these small initial perturbations are the seeds of today's structure, and their growth and evolution is governed by gravitational collapse (Mukhanov, 2005). The role of gravity is to drive the collapse of overdensities, which in turn results in the emptying out of under-densities. As these density fluctuations grow, their evolution can be initially described in a perturbative regime, known as the regime of linear growth where $\delta << 1$. In this linear regime the growth of perturbations can be expressed via a multiplicative *linear growth function* $D(a)$ such that at any spatial length scale $\delta(a) \propto D(a)$ (Huterer et al., 2015).

A useful description of the properties of the density field is done through the 2-point correlation function $\xi(r) = \left\langle \delta(\mathbf{x})\,\delta(\mathbf{x}') \right\rangle$ with $r = \left|\mathbf{x} - \mathbf{x}'\right|$ and the power spectrum $P(k)$ which is the Fourier transform of the correlation function. A commonly used parameter relating to this is the amplitude of density fluctuations σ_R measured at smoothing scale R, which can be expressed from the (linear) power spectrum as

$$\sigma_R^2(a) = \int_0^\infty \frac{k^3}{2\pi^2}\, P_{\rm lin}(k,\, a)\, W^2(kR)\, \mathrm{d}ln\, k\,, \tag{1.15}$$

where W is the Fourier transform of the spherical top hat filter (Huterer et al., 2015). The value of $\sigma_8 = \sigma_R(a = 1)$ with $R = 8\,\mathrm{Mpc}\,h^{-1}$[4] is a commonly used cosmological parameter and its value is measured to be $\sigma_8 \approx 0.81$ (Planck Collaboration et al., 2018).

As perturbations grow, eventually they reach the non-linear phase when $\delta \sim 1$ or greater. An illuminating example for this phase is the gravitational collapse of an overdensity distributed

[3] Alternatively, in an inflationary picture the process of inflation drives K to approach zero(Mukhanov, 2005)

[4] $h = H_0/(100\,\mathrm{km/s/Mpc})$

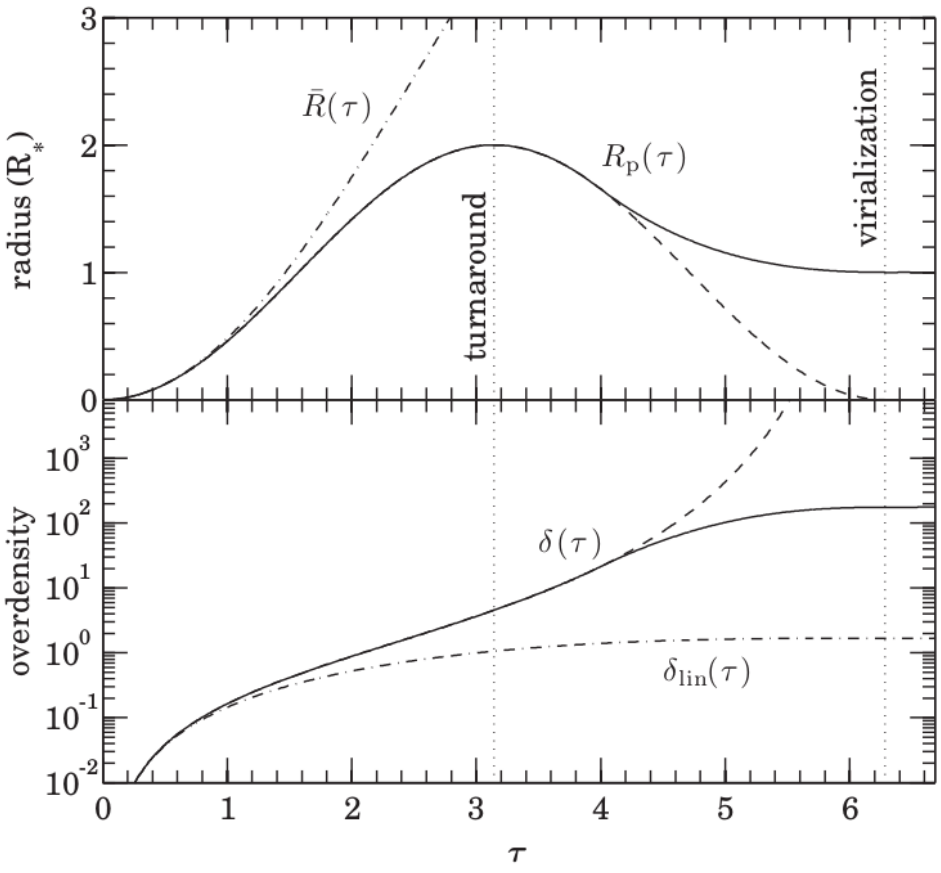

Figure 1.1: Illustration of the spherical collapse model. *top:* Real space expansion and collapse. *bottom:* change in density contrast according to linear and non-linear theory. Figure adapted from Knobel (2012).

according to a spherical top hat embedded in an expanding homogeneous background. An illustration of this collapse is shown on top panel of Figure 1.1: Initially each concentric mass shell of the overdensity follows the background expansion or *Hubble flow* indicated as $\bar{R}(\tau)$, but over time they decouple from it indicated by $R_P(\tau)$ and start contracting under their self-gravity. A perfectly spherical top hat overdensity would collapse into a point. For more general or perturbed distributions instead the system reaches *virial equilibrium* where the kinetic energy of the dissipationless system counteracts further gravitational collapse. In this final virialized stage the system contracted to approximately half its size at turnaround[5] (Kravtsov & Borgani, 2012; Knobel, 2012). Hence this collapse process out of an initial small overdensity creates a gravitationally bound object known as a *halo*.

The bottom panel of Figure 1.1 shows the above spherical collapse process, but expressed in terms of the density contrast of the collapsing volume $\delta(\tau)$ and the density contrast as predicted by the growth of linear perturbations $\delta_{\rm lin}(\tau)$. Initially these two descriptions agree (when $\delta << 1$), however as the collapse proceeds these two descriptions deviate. While linear theory predicts $\delta_c = 1.69$ at the time of the collapse, the actual virial overdensity is $\Delta_{\rm vir} = \delta_{\rm vir} + 1 \approx 178$ independently of the mass of the formed halo. Hence halo masses are commonly quoted as the mass enclosed within a sphere having the virial overdensity $M_{\rm vir}$ or 200 times the cosmic average matter density M_{200m}.

Linear theory nominally breaks down for $\delta \sim 1$, yet the above finding still means that overdensities for which $\delta_{\rm lim} > \delta_c = 1.69$ will form virialized halos. This connection enables us to express the number and mass distribution of said halos from linear theory, defining the *halo mass function* (Huterer et al., 2015)

$$\frac{{\rm d}n}{{\rm d}M} \equiv \frac{{\rm d}^2N}{{\rm d}M\,{\rm d}V} = F(\sigma)\,\frac{\rho_M}{M}\frac{{\rm d}\ln\sigma^{-1}}{{\rm d}M} \tag{1.16}$$

where σ is defined according to Equation 1.15 but using the the radial scale corresponding to $M = (4\pi/3)\,R^3\rho_M$ with $\rho_M = \Omega_M\,\rho_{\rm crit}$, and $F(\sigma)$ denotes the fraction of collapsed objects, which according to the formalism of Press & Schechter (1974) can be estimated as

$$F(\sigma) = \frac{1}{\sqrt{2\,\pi\,\sigma^2}}\exp\left(-\frac{1}{2}\frac{\delta_c^2}{\sigma^2}\right)\,. \tag{1.17}$$

Equation 1.16 makes it explicitly visible that the halo mass function captures the abundance of halos both in terms of volume density and also in terms of density per mass element. Hence halos provide a cosmological probe which is sensitive to both the geometry of the Universe, and also to the growth of cosmic structure (Huterer & Shafer, 2018). While the above formulas are an elegant way of expressing the halo mass function from linear theory, it is not sufficiently accurate for the needs of precision cosmology. For this reason studies in recent literature adopted empirical formulas motivated by the above, which are fitted to the measured halo mass functions in numerical simulations (Tinker et al., 2008; McClintock et al., 2019).

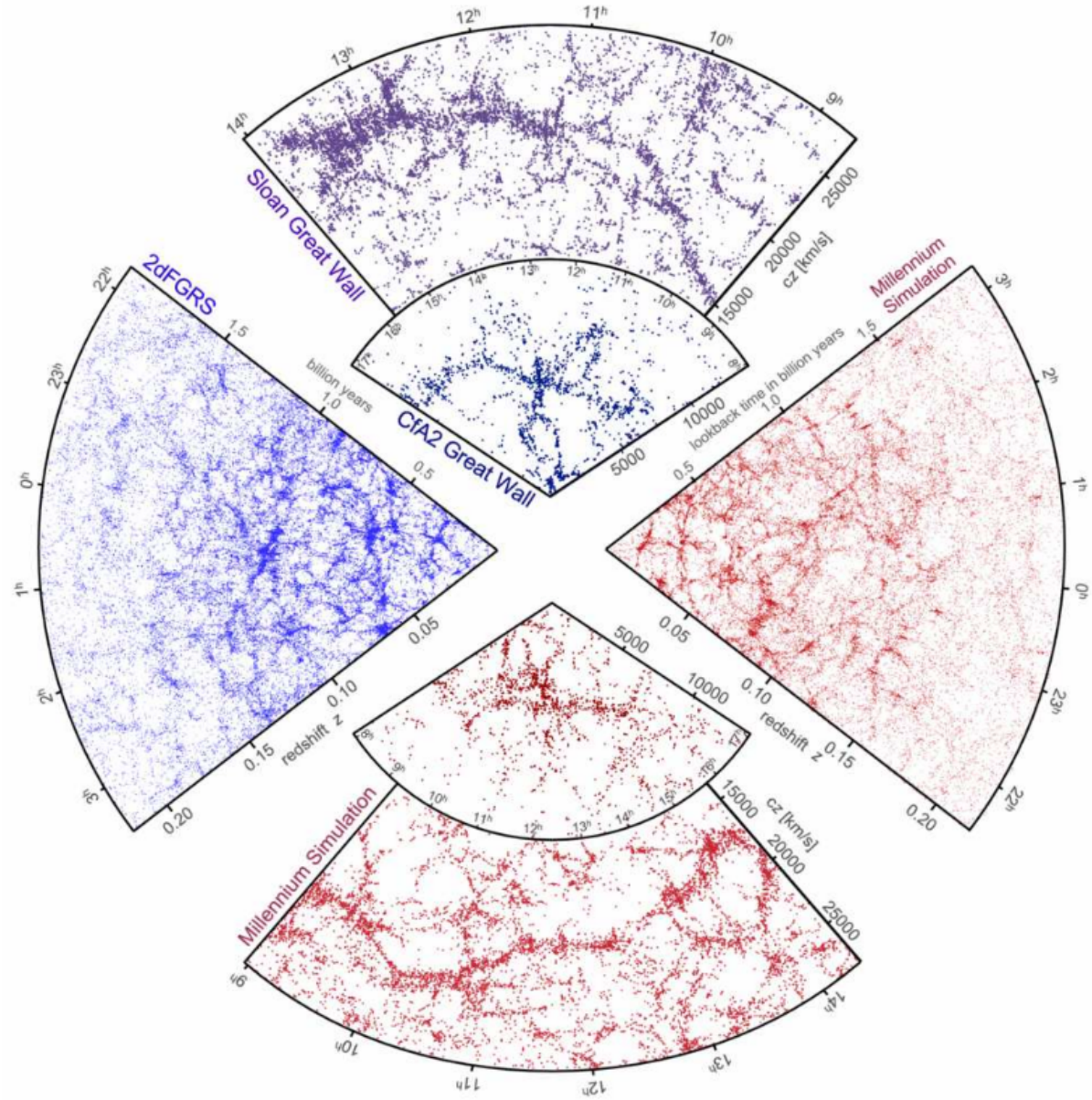

Figure 1.2: The large scale structure of the Universe in observations as traced by galaxies (left and upper segments) and in numerical simulations as traced by dark matter halos (bottom and right segments). The radial direction represents the redshift direction. Figure adapted from Springel et al. (2006).

1.1.4 The Cosmic Web

The gravitational collapse of initial overdensities in the cosmic matter field gives rise to bound objects known as halos, but the time it takes to form a virialized halo depends on the spatial extent of the overdensity: Smaller spatial scales undergo collapse, decoupling and virialization quicker than larger scales. The size and density of the formed halo is relative to the cosmic matter density at the time of its formation, therefore in an expanding cosmic background halos which form earlier are not only smaller but also denser than their later forming counterparts. Furthermore, due to the disparate collapse timescales involved, some smaller halos are already fully formed by the time larger scale overdensities start to collapse, therefore these superstructures form out of the *mergers* of smaller halos (Springel et al., 2006). This phenomenon is often denoted as the hierarchical formation of cosmic structure, and results in a system where halos are not simply smooth overdensities but themselves are composed of smaller *subhalos* embedded in the superstructure.

Based on the local geometry of overdensities, gravitational collapse can happen not only in three dimensions, which would form halos, but also in two dimensions, forming structures known as *filaments*[6]. This type of process results in a pattern known as the *cosmic web*. Because of the hierarchical formation smaller halos are already embedded in the currently collapsing structures and can act as tracers of the large scale structure. Such a scenario is illustrated on Figure 1.2, which shows the cosmic web made visible through the distribution of galaxies in different observational datasets (left and upper segments) and in numerical simulations as the distribution of halos (right and bottom segments) (Springel et al., 2006). Both observational data and simulation results paint a consistent picture where filaments of matter, halos and galaxies connect into the knots of the cosmic web. Meanwhile initial underdensities grow emptier as time passes, visible only as *voids* bounded by overdense filaments and knots.

1.2 Non-linear Baryonic Structures

In a direct observational scenario, of the above described structures we can only see the baryonic matter component. In particular, the visible segment of the electromagnetic spectrum is dominated by stellar emission, and for this reason the galaxies in which these stars reside became a fundamental tracer of cosmic structure. Galaxies initially form within their own halo, however through mergers over cosmic history some of them end up embedded in larger structures known as *galaxy groups* and *galaxy clusters* (Kravtsov & Borgani, 2012), and became subhalos within them.

1.2.1 Galaxies in Optical Wavelengths

Galaxies are gravitationally bound systems of stars, gas and dust within a dark matter halo or subhalo. While their mass is dominantly in the form of dark matter, their most apparent observa-

[5] For a virialized system the total kinetic energy K and the potential energy V is related as $K = -1/2V$.

[6] and also in 1D forming so-called sheets, however these are less manifest in simulations or observations.

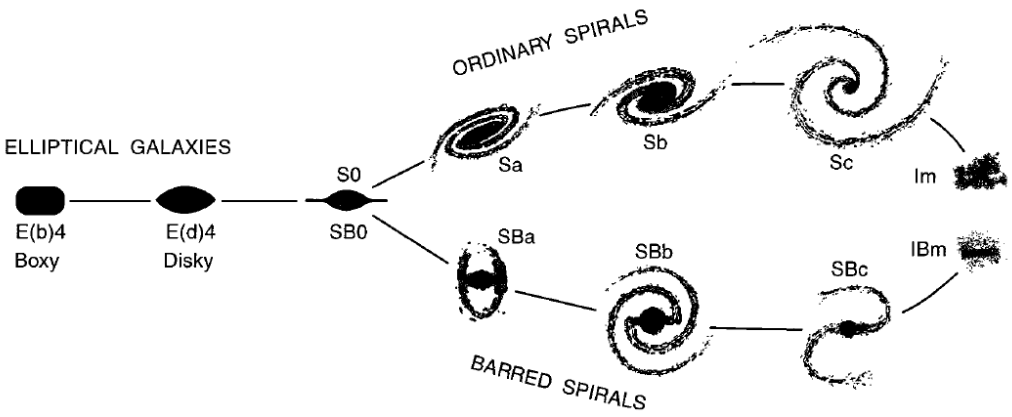

Figure 1.3: Morphological classification of galaxies along the Hubble Sequence. Figure adapted from Kormendy & Bender (1996).

tional feature is that they are the environments where star formation can efficiently happen within the cosmos. Stars are formed out of the gravitational collapse of gas clouds, and thus sustained star formation requires both a steady gas reservoir and a process for gas to cool and contract by radiating its thermal energy away.

The visible properties of galaxies are closely connected to the properties of their current stellar components, and this in turn is the product of their *star formation history* over cosmic time. The sum of stellar light determines the net *spectral energy distribution* (SED), while the orbital distribution of stars within the potential well determines the *morphology* of the galaxy. From the morphological point of view galaxies are traditionally classified along the Hubble sequence illustrated in Figure 1.3, where they are grouped into two main groups based on their light distribution: elliptical and spiral galaxies (Kormendy & Bender, 1996).

Spiral galaxies are named after the prominent spiral arm features visible in them which are regions of recent star formation. They are also characterized by a galactic disk defined by the largely coherent orbits of stars and gas within the potential well, and a central stellar component known as the bulge which is classically regarded as dynamically distinct from the disk. In some cases spiral galaxies also possess a bar-like feature defined by the stellar orbits in their central region which leads to further morphological and dynamical classifications. Elliptical galaxies in turn are characterized by the chaotic individual orbits of stars and it is their velocity dispersion which keeps the system in gravitational equilibrium. They are thus described by a simpler set of morphologies as triaxial ellipsoids, which is often expressed as the shapes of their isophotes. There are several narrower categories beyond the top-level simple picture including S0 galaxies inhabiting a transitional morphology between spirals and ellipticals, and different types of irregular galaxies which do not fit into other categories.

The typical spectral energy distribution of different morphological types is illustrated in Figure 1.4 (Molino et al., 2014). These SEDs form a sequence from elliptical (Ell) galaxies, through spirals (Sbc, Sbd) to starburst galaxies (SB). This sequence is defined by the age of the stellar population of the galaxy, therefore it is visible that spiral and elliptical galaxies differ not only

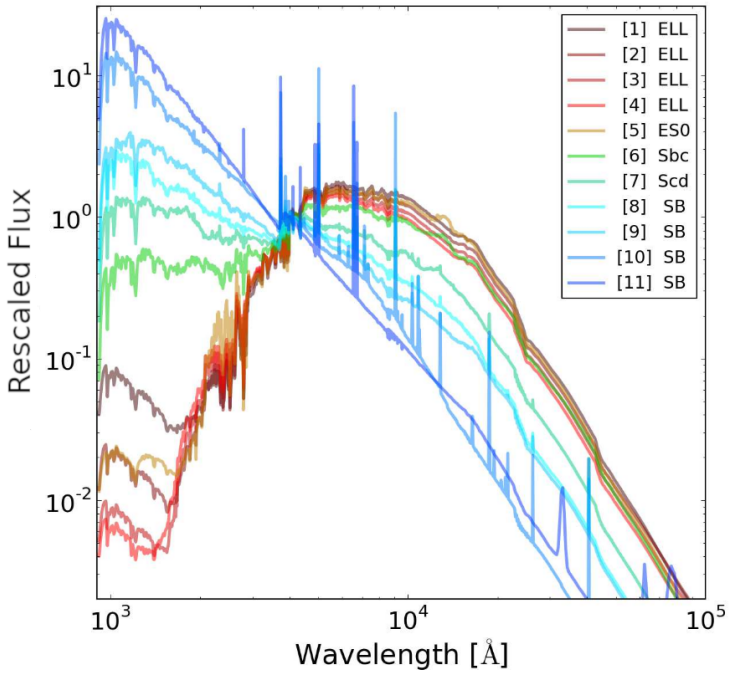

Figure 1.4: Typical galaxy spectral energy distributions (SEDs) for different morphological types and starburst galaxies (SB). Figure adapted from (Molino et al., 2014).

in their morphology but also in their star formation history. Elliptical galaxies have a very old, passively evolving stellar population, while spiral galaxies have ongoing star formation and their overall stellar components are younger. The SEDs of star forming galaxies are dominated by the freshly formed hot young stars. When star formation ceases, an event known as *quenching*, the short lived hot stars expire soon after[7], and only colder and redder stars remain. As quenched galaxies age, their SEDs change only through the passive evolution of their stars. For actively star forming galaxies however the SED is strongly determined by the rate and history of star formation, which results in a substantially greater variability within and between the different galaxy types.

A practical way of describing galaxies in observational studies is by positioning them on a color-magnitude diagram. This carries similar information to comparing the SEDs of different galaxies, and measuring magnitudes and colors (magnitude differences) is significantly easier than measuring spectra. Figure 1.5 shows the color-magnitude diagram of an exemplary galaxy population, where in addition the morphological classification of the galaxies is also shown. Ellipticals are indicated by red markers, while different types of spiral galaxies are shown as blue or green markers. It is visible that the galaxies are divided into two populations, a preferentially bluer spiral population distributed in the *blue cloud*, and a redder elliptical galaxy population distributed along the *red sequence* (Tinker et al., 2013). Galaxies can become brighter by form-

[7]Stellar life-cycles of 10-100 million years for O or B type stars are almost instantaneous compared to cosmological timescales

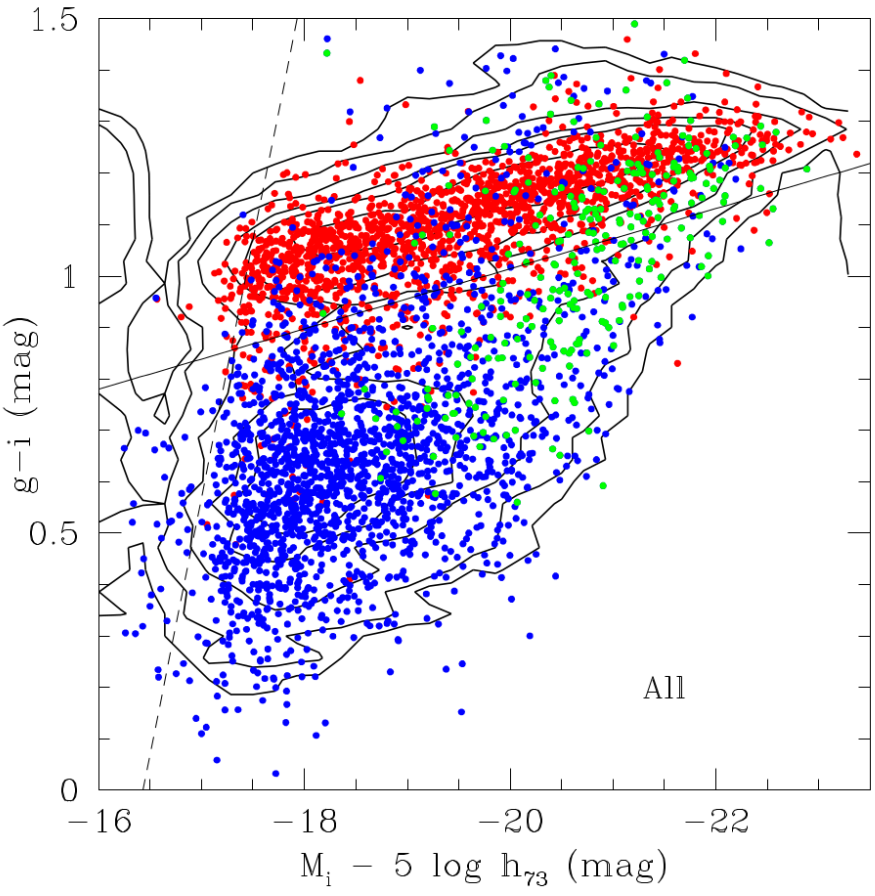

Figure 1.5: Color-magnitude diagram of galaxies in the Sloan Great Wall (see uppermost corner segment of Figure 1.2). The marker color indicates observed morphology *red:* elliptical and S0, *green:* Sa and Sb, *blue:* generic spirals and irregulars. Figure adapted from Gavazzi et al. (2010).

ing stars, however the merger of galaxies is also a possible pathway of growth. Mergers disturb the dynamical state of galaxies, disrupt the disk structure of spiral galaxies, and they can induce quenching. This is indeed visible on Figure 1.5 as the brightest galaxies are preferentially red ellipticals, which were able to increase their brightness by incorporating the the stellar components of other galaxies via mergers as opposed to in-place star formation.

1.2.2 Galaxy Clusters

The most massive gravitationally bound structures in the Universe today are galaxy clusters, which correspond to the highest peaks in the initial cosmic density field. These massive halos form out of the mergers of smaller galaxy groups and galaxies in a hierarchical way (Kravtsov & Borgani, 2012; Allen et al., 2011). The distinction between galaxy groups and clusters is somewhat arbitrary, both being the product of the same formation process, albeit the terms indicate different mass scales with clusters commonly denoted as halos with masses of $M_{200} \geq 10^{14}$ $M_{\odot}$. Mergers promote quenching in the infalling galaxies which leads to galaxy clusters hosting a characteristic population of quenched elliptical galaxies distributed along the red-sequence. Thus the bulk of the stellar mass within a cluster is actually formed outside of the halo, when the eventual cluster member galaxies were still independent and the dominant members of their own halos.

A prominent element of galaxy groups and clusters is the central galaxy (or brightest cluster galaxy, BCG) which in an equilibrium system is positioned at the deepest point of the potential well[8], while the other member galaxies inhabit subhalos and orbit within the main cluster halo. In this setting strong tidal effects act on the subhalos, stripping matter from them and transferring it to the main cluster halo. The stellar component of this stripped material is visible as the intracluster light (ICL), which together with the central galaxy is estimated to contain a substantial fraction of the stellar mass of the galaxy cluster (Kluge et al., 2020; Zhang et al., 2019a). An often used metric of the galaxy content of clusters is the *richness* which is defined as the number of galaxies in the cluster brighter than some limiting magnitude, although in practice richness estimates are usually formulated with a more narrow galaxy selection in mind. Figure 1.6 shows the optical g, r, i band color composite image of the galaxy cluster MACS J0416.1-2403, which is one the most massive known halos identified to date, visible here as an overdensity of luminous red-sequence galaxies (Ebeling et al., 2001).

The luminous stellar component constitutes only approximately 1% of the total mass of a galaxy cluster, with most of the matter being in the dark matter halo, and approximately 10% of the mass is present in the form of gas bound in the cluste rs gravitational potential, also known as the intra-cluster medium. Due to the depth of the potential well, accreted gas infalling to the cluster is heated to very high temperatures (up to 10^8 K) and radiates thermally in X-ray wavelengths (Kravtsov & Borgani, 2012). The hot gas (plasma) component gives rise to an additional pathway for the direct observation of galaxy clusters: Photons originating from the cosmic microwave background (CMB) suffer inverse Compton scattering on the electrons of

[8]During mergers of similarly sized clusters it can sometimes appear that that there are more than one similarly bright galaxies, or that the brighter galaxy is in fact not the one closer to the center of the mass distribution.

the plasma in the ICM, a process known as the Sunyaev-Zel'dovich effect. The result of this scattering is that depending on the observed wavelength range clusters appear as shadows or bright spots on maps of the CMB.

The dark matter, gas and galaxy components of galaxy clusters show a large degree of self-similarity in their distributions across a wide range of masses. This in practical terms means that clusters at different masses can be well described as scaled versions of each other (Kravtsov & Borgani, 2012). Self-similarity is also manifested in a universal density profile for cluster sized halos, the Navarro-Frenk-White NFW profile (Navarro et al., 1996)

$$\rho_{\mathrm{NFW}}(r) = \frac{\rho_{\mathrm{crit}}\,\delta_c}{\left(\frac{r}{r_s}\right)\left(1+\frac{r}{r_s}\right)^2}\,, \tag{1.18}$$

where r_s is the scale radius of the halo, while δ_c is known as the characteristic overdensity of the halo. In practice these parameters are more commonly expressed equivalently by the radius containing 200 or 500 times the cosmic mean matter density r_{200m} or r_{500m}, and the concentration parameter c respectively.

1.3 Weak Gravitational Lensing

In the language of general relativity, light rays follow the null geodesics of spacetime where $ds^2 = 0$. In the presence of massive objects these paths are no longer straight lines, and light rays become deflected due to the inhomogeneous gravitational potential. This effect is called *gravitational lensing* (Bartelmann & Schneider, 2001).

In the following we introduce a specific lensing scenario following Bartelmann & Schneider (2001), illustrated in Figure 1.7, which is applicable for the weak field limit of general relativity. Here a light bundle emitted by a distant source is deflected by the gravitational potential of a lens mass distribution before it reaches the observer. For small deflection angles the trajectories of light rays can be approximated as straight lines before and after the deflection. Furthermore in the cosmological context the relevant distances are much larger than the size of the lens, motivating the thin-lens approximation where the lens mass distribution can be expressed as a surface mass density $\Sigma(\boldsymbol{\xi})$ localized on the lens plane. For this setting the deflection angle is expressed as

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2}\int \mathrm{d}\xi^2 \Sigma(\boldsymbol{\xi}')\frac{\boldsymbol{\xi}-\boldsymbol{\xi}'}{\left|\boldsymbol{\xi}-\boldsymbol{\xi}'\right|^2}\,. \tag{1.19}$$

The true angular position of a light bundle or *source* at which it would be visible to the observer without lensing β, and the lensed or deflected angular position θ, are connected via the *lens equation*

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{ds}}{D_s}\hat{\boldsymbol{\alpha}}\left(D_d\boldsymbol{\theta}\right)\,, \tag{1.20}$$

where D_s, D_d, D_{ds} refer to angular diameter distances (Equation 1.13) to the source, to the lens, and between the source and the lens respectively.

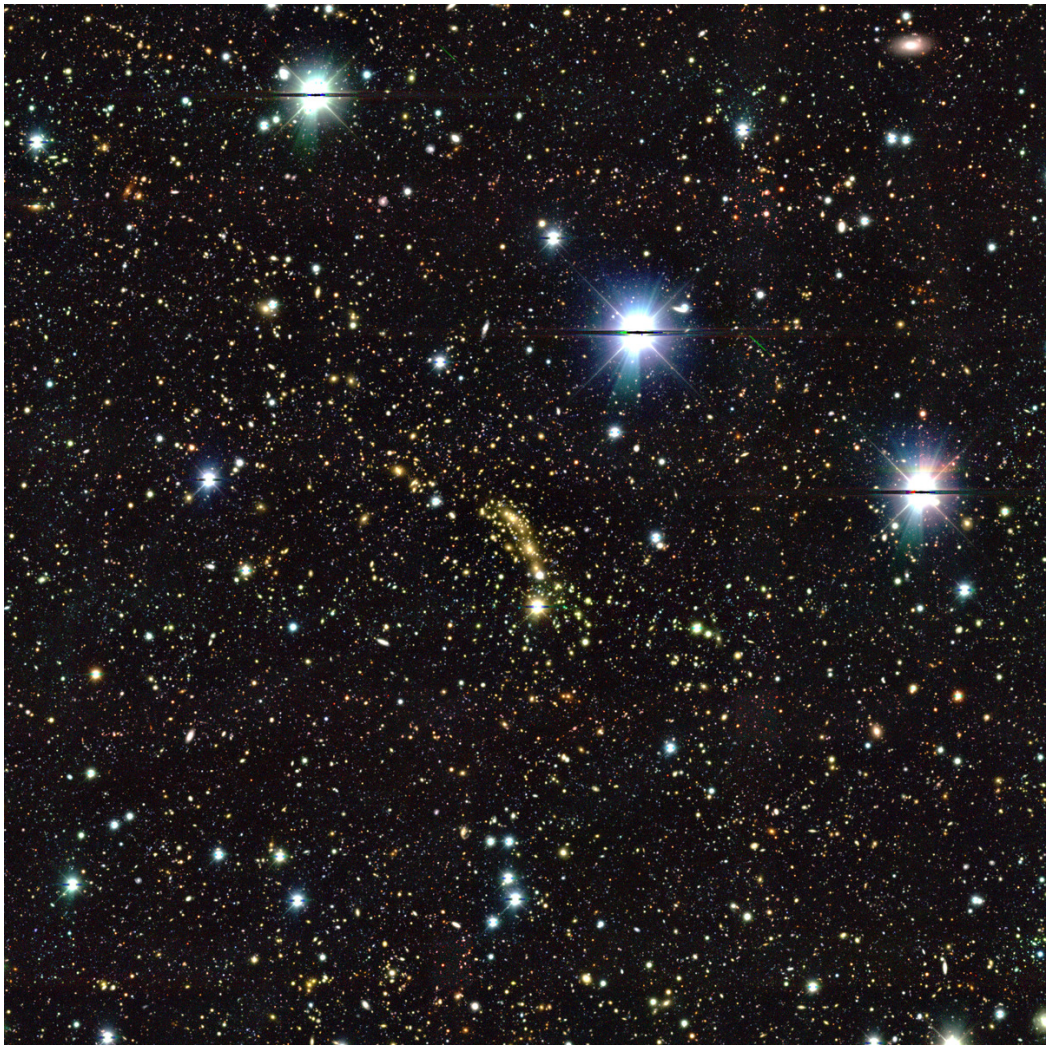

Figure 1.6: Color composite image (g, r, i) of the galaxy cluster MACS J0416.1-2403. Figure courtesy of Daniel Gruen and the DES deep field program.

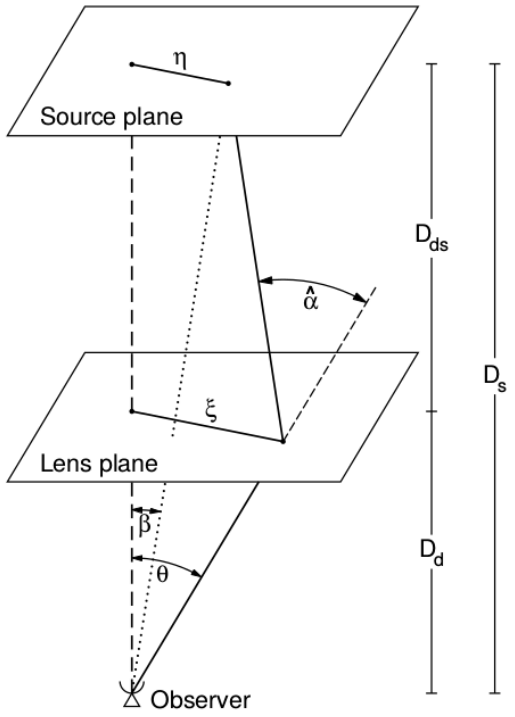

Figure 1.7: Lens geometry and light deflection path in a simple weak lensing scenario. Figure adapted from Bartelmann & Schneider (2001).

Depending on the mass distribution and the distances involved, the lens equation can have multiple solutions for a single $\boldsymbol{\beta}$, known as multiple images. The criterion for this is given by the *critical surface mass density*

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_s}{D_d\, D_{ds}}\,, \tag{1.21}$$

which is used to define the convergence or dimensionless mass density

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(D_d\boldsymbol{\theta})}{\Sigma_{\rm crit}}\,. \tag{1.22}$$

$\kappa \geq 1$ is a sufficient criterion to produce multiple images and this is known as the *strong lensing* regime, while $\kappa < 1$ is known as the *weak lensing regime*, although this term is most often used for $\kappa << 1$.

During lensing light bundles suffer differential deflection, which results in the change in their size and shape. This mapping can be locally linearised and it is then described by the weak lensing Jacobian matrix

$$\mathcal{A}(\boldsymbol{\theta}) = \frac{\partial\boldsymbol{\beta}}{\partial\boldsymbol{\theta}} = \left(\delta_{ij} - \frac{\partial^2\Psi(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right) = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}\,, \tag{1.23}$$

where Ψ is the *lensing potential* defined from the convergence in a 2D analogy of the classical gravitational potential,

$$\Psi(\boldsymbol{\theta}) = \frac{1}{\pi}\int \mathrm{d}^2\theta'\,\kappa(\boldsymbol{\theta}')\ln\left|\boldsymbol{\theta}-\boldsymbol{\theta}'\right|\,, \tag{1.24}$$

and a new quantity: *gravitational shear* $\boldsymbol{\gamma} \equiv \gamma_1 + i\gamma_2$ is introduced. According to the lensing Jacobian, initially circular light sources are mapped into ellipses. The anisotropic shape distortion is described by the shear, while the change in area is releated to both κ and γ as expressed by the *magnification*

$$\mu \equiv \frac{1}{\det A} = \frac{1}{(1-\kappa)^2-\gamma^2}\,. \tag{1.25}$$

The great practical advantage of weak lensing relies in the fact that by observing the distorted images we can reconstruct some properties of the lens potential. To demonstrate this, first let us consider that observationally one does not simply measure $\boldsymbol{\gamma}$, but the *reduced shear*

$$\boldsymbol{g} = \frac{\boldsymbol{\gamma}}{1-\kappa}\,, \tag{1.26}$$

which is directly related to the *ellipticity* $\mathbf{e} = e_1 + ie_2$ of the distorted light bundle defined such that $|e| = (a-b)/(a+b)$ where a and b are the major and minor axes of the light bundle.[9]

So far we invoked the assumption of initially circular light bundles, but lack of access to such idealized tracers does not limit the applicability of weak lensing measurements. In a general case we can consider lensing experienced by an ensemble of tracers, such as background *source galaxies* each with an intrinsic shape and size. The shape of source galaxies can be characterized by the ellipticity $\mathbf{e}$ or second moment of their light distributions. It is assumed that true background galaxies have no preferential intrinsic orientation $\langle \mathbf{e}_{\rm intr}\rangle \approx 0$, thus in an unbiased measurement scenario the reduced shear experienced by an ensemble of tracers can be expressed as $\langle \mathbf{e}\rangle \approx \langle g\rangle$. This naturally depends dominantly on the accuracy of estimating galaxy shapes, and only holds exactly in an ideal scenario.

Gravitational shear is directly related to the mass distribution: Each mass element in the lens plane induces a shear field which is aligned tangentially with respect to it, and the net shear at each point is found as the integral over all mass elements (Kaiser & Squires, 1993). In turn, the net tangential shear measured at projected radius R around a lens can be expressed as

$$\langle \gamma_{\rm T}(R)\rangle = \overline{\kappa}(<R) - \langle \kappa(R)\rangle\,, \tag{1.27}$$

where $\langle\cdot\rangle$ refers to the expectation value integrated over a circle of radius R, and $\overline{\kappa}(<R)$ refers to the mean convergence measured within the disc of radius R around the targeted point (Mellier, 1999). This equation can be transformed into the actual surface mass density by multiplying with $\Sigma_{\rm crit}$

$$\Sigma_{\rm crit}\cdot\langle \gamma_{\rm T}(R)\rangle = \overline{\Sigma}(<R) - \langle \Sigma(R)\rangle \equiv \Delta\Sigma(R) \tag{1.28}$$

where $\Delta\Sigma$ is the *excess surface mass density*. This equation is formulated for $\gamma_{\rm T}$, and is therefore only fully applicable for measurements of $g_{\rm T}$ in the weak shear limit. Furthermore this type of

[9] In the weak shear limit of $\gamma \ll 1$ and $\kappa \ll 1$, $g \approx \gamma$.

inversion constrains the surface mass density only up to an additive constant, a phenomenon known as the mass sheet degeneracy. Nevertheless, knowing a theoretical model for the lens mass distribution, $\Delta\Sigma$ can be predicted, and thus the total mass of the lens system can be derived from the model fit. This is the method which serves as the basis for the weak lensing mass calibration presented in this thesis.

1.4 The Dark Energy Survey

The Dark Energy Survey (DES) is an optical imaging sky survey covering a footprint of 5000 deg^2 near the southern galactic cap with g, r, i, z band observations (DES Collaboration et al., 2016). The survey is conducted using the 570 megapixel Dark Energy Camera (DECam) installed on the 4-m Blanco Telescope at the Cerro-Tololo Inter-American Observatory (CTIO) (Flaugher et al., 2015). The primary science driver of DES is to constrain the dark energy equation of state parameter w (see Equation 1.6) from a combination of cosmological probes sensitive to the growth of structure and the expansion of the Universe at late cosmic times. In this context the design and operations of DES are centered around using weak gravitational lensing to connect the properties of directly observable tracers such as galaxies or galaxy clusters with the underlying distribution of dark matter.

Observations in DES are organized into six seasons. These were preceded by a science verification (SV) survey performed from November 2012 to February 2013 at a restricted footprint of approximately 140 deg^2 but at the eventual target depth of DES corresponding to a 10σ i-band limiting magnitude of 23.5 (DES Collaboration et al., 2016). The main survey operations started in August 2013, and these consist of three main data and science releases based on the data collected up the first year (Y1), third year (Y3) and fifth year (Y6) of observations. Currently DES is analyzing data taken up to February 2016 at the end of the third year of observations. The cumulative footprints of the different seasons are shown on Figure 1.8 where the Y1 area is indicated in red, while the black contour shows the footprint of the Y3 and Y6 areas. In addition to the main survey footprint DES also covers a set of deep field regions aimed at supporting the DES supernova cosmology program (Abbott et al., 2019) by repeated observations at an approximately weekly cadence, while also complementing the DES wide field survey with approximately 2-3 mag deeper observations.

One of the primary data products of the DES main survey is a base photometry catalog (denoted as the GOLD catalog) derived from the galaxies and stars identified in the survey footprint of an observational season. These catalogs, i.e. the Y1 GOLD (Drlica-Wagner et al., 2018) or Y3 GOLD (DES Collaboration, in prep a) catalogs are the primary science ready photometric data products of DES. In addition to the fluxes and basic photometric properties of each object, they contain the photometric calibrations and corrections and ancillary data products such as observational masks or star-galaxy separation flags needed to define a suitable input dataset for cosmology analyses.

While the relevant chapters of this thesis introduce the relevant data products and concepts in more detail, within the following sections we recapitulate some of the general methods and considerations underlying the thesis.

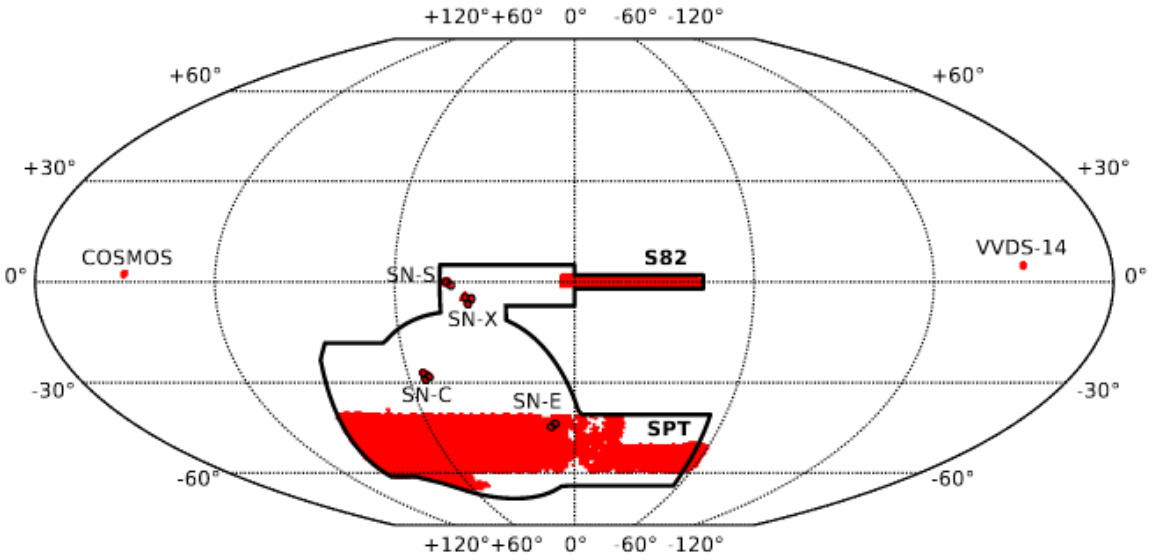

Figure 1.8: Observational footprint of the Dark Energy survey. *(red area:)* year 1 (Y1) footprint. *(black circles:)* Deep field regions. *(black contour:)* year 3 (Y3) and year 5 (Y6) footprint. Figure adapted from Drlica-Wagner et al. (2018)

1.4.1 Cluster Finding in Photometric Data

In optical wavelengths galaxy clusters appear as overdensities of galaxies. Many of these cluster member galaxies are quenched, elliptical galaxies distributed along the red sequence. While not all cluster members are red or elliptical, in the context of identifying clusters, searching for them as overdensities of red-sequence galaxies presents a powerful and fundamental method, as their SED and evolution can be modeled more simply than that of other galaxy types (Koester et al., 2007; Weinberg et al., 2013). Furthermore this kind of approach benefits from the fact that it can directly be applied in the photometric catalogs of wide area sky survey such as DES at no additional observational cost.

Within DES the galaxy clusters are identified using the redMaPPer: **red**-sequence **Ma**tched-filter **P**robabilistic **Per**colation algorithm (Rykoff et al., 2014, 2016). The base input for redMaPPer is the DES GOLD photometry catalog and a small set of spectroscopic seed galaxies. These seed galaxies are used to train the initial model for the red-sequence at different redshifts which is then iteratively expanded and re-trained with the galaxies in the photometric dataset. The algorithm uses a spatial and color based matched-filter to identify overdensities of galaxies which are likely located at the same redshift, and from these constructs catalogs of cluster candidates. Cluster candidates are assumed to be centered on one of the galaxies in the photometry catalog, while galaxies which according to the matched-filter might be red members of a cluster candidate become possible members with a membership probability p_{memb}.

RedMaPPer derives cluster redshifts from the p_{memb} weighted individual red-sequence based photometric redshifts of potential member galaxies, and assigns an optical richness estimate λ to

each cluster candidate which is defined based on the effective number of cluster members as

$$\lambda = \sum_i p_{\mathrm{memb};\,i} \;:\; R_i < R_\lambda \text{ and } L/L_\star > 0.2\,. \tag{1.29}$$

Here R_i is the projected radius from the cluster center expressed in physical units at the redshift of the cluster candidate, while R_λ is a search radius chosen maximize the correlation of lambda with cluster mass. $L/L_\star$ is the relative luminosity of the potential member galaxy compared to the characteristic galaxy luminosity at the considered redshift.

Based only on the matched-filtering step, the same cluster might be found multiple times with slightly different central galaxy candidates, and the separation of structures which overlap in projection is ambiguous. This is remedied by the percolation component of redMaPPer where in an iterative approach member galaxies are reassigned and the p_{memb} values are updated, and clusters candidates are joined or separated until convergence is reached with respect to the richness and redshift values of the surviving cluster candidates, thus defining the redMaPPer cluster catalog. In this percolation step member galaxies are re-evaluated as a potential cluster centers, and based on this the final cluster center is determined in a probabilistic way with potentially multiple centeral galaxies each with probability p_{cen}.

In practice studies using redMaPPer employed two distinct type of cluster catalog: a flux-limited, and a locally volume limited catalog. The difference between these two versions is their treatment of the survey limiting magnitude. The volume limited catalog considers only clusters where the galaxy luminosity threshold in Equation 1.29 is above the survey magnitude limit, effectively defining a redshift or volume limit for the catalog. Given that the survey limiting magnitude can vary within the footprint due to the unevenness of observational conditions or gaps this is referred to as a *locally* volume limited selection. In turn, the flux limited version considers cluster candidates where some of these galaxies are undetected, and their presumed contribution to the richness of the cluster is extrapolated based on the brighter detected cluster members. In principle this allows greater redshift range for the cluster catalog, however at the cost of increased richness and redshift uncertainty.

Complementing the galaxy cluster catalog, redMaPPer also creates a catalog of reference random points, whose spatial and redshift distribution is representative of the footprint and volume of the cluster catalogs. The shape of the footprint for a sample of galaxy clusters depend on both their richness and redshift. This behavior originates from the gaps and non-trivial masking geometry of the survey, which means that a low richness higher redshift cluster may fit into a given unobstructed area in the survey, while a lower redshift higher richness cluster may have a large fraction of its galaxies masked from the input catalog. For this reason random points are labeled also with a richness and redshift value, using which, one can reconstruct the actual survey volume and footprint of the specific type of clusters.

1.4.2 Galaxy Shape Estimates

As introduced in section 1.3, the excess surface mass density profile of a gravitational lens can be inferred from the preferential tangential alignment of background source galaxies. This measurement relies on estimating the redshifts of the source galaxies, that is estimating the lens-geometry,

and on estimating the shapes and alignments of the images of source galaxies. In the following we briefly review the practical methods and considerations involved in performing galaxy shape measurements in the Dark Energy Survey based on Jarvis et al. (2016) and Zuntz et al. (2018).

1.4.2.1 Data Preparation

The aim of this effort is to construct a catalog of source galaxies where for each an ellipticity (shape) value $\mathrm{e} = e_1 + i\,e_2$ is estimated, along with the necessary corrections and photometric selection flags. The starting dataset for this are the single exposures taken by DES in the different observing bands, and data reduction is done in parallel with the construction of the GOLD photometry catalog. Image processing starts with low-level calibration of single-band single-exposure images to detect observational artifacts, measure the noise levels and regularize the observations. Following this, single exposure images are coadded into deeper detection images and the GOLD photometry catalog is created (Drlica-Wagner et al., 2018). At this stage stars and galaxies are distinguished based on their different distributions on a size–magnitude diagram, and a subset of suitable stars are selected to estimate and model the *point spread function* (PSF) on each single exposure[10].

The next step in data processing is to collate the observational data for each photometry detection listed in the GOLD catalog. This is done by creating small cut-out images or *postage stamps* of each object in all single exposure images and in each band, and these are collected into a multi-epoch data structure (MEDS). This way the available observational data can be queried efficiently based on the selected object, while treating different exposures with different PSFs and observational conditions separately.

1.4.2.2 Galaxy Shapes and Observational Effects

Before discussing the actual shape measurement methods employed in DES, it is illuminating to consider how the observed signal comes to be. This is illustrated on Figure 1.9, where the observed images on the rightmost panels roughly correspond to what is visible on a postage stamp image, and where the top and bottom rows correspond to galaxies and stars respectively. Galaxies are extended objects with potentially complicated morphologies, however in a practical scenario we have little a priori information about their true individual intrinsic properties such as their sizes, shapes or morphological types. Throughout the journey to the observer, a gravitational lensing signal g is imprinted on the images of galaxies which induces a small distortion in them (exaggerated on Figure 1.9). The quantity of interest when constructing a weak lensing source catalog is this lensed shape or ellipticity, however this cannot be directly measured. The actual detected images are produced as a product of additional observational and instrumental effects. These include a convolution (smearing) by the PSF of the atmosphere and the telescope, a discretization of the signal due to the pixel grid of the detector, and finally noise originating from the detector and atmospheric background.

While the above series of observational effects imprinted on galaxy images present a daunting scenario, in fact they can be measured and characterized directly from observational data.

[10] In the case of DES Y1 in fact the PSF is estimated and modeled for each CCD detector separately

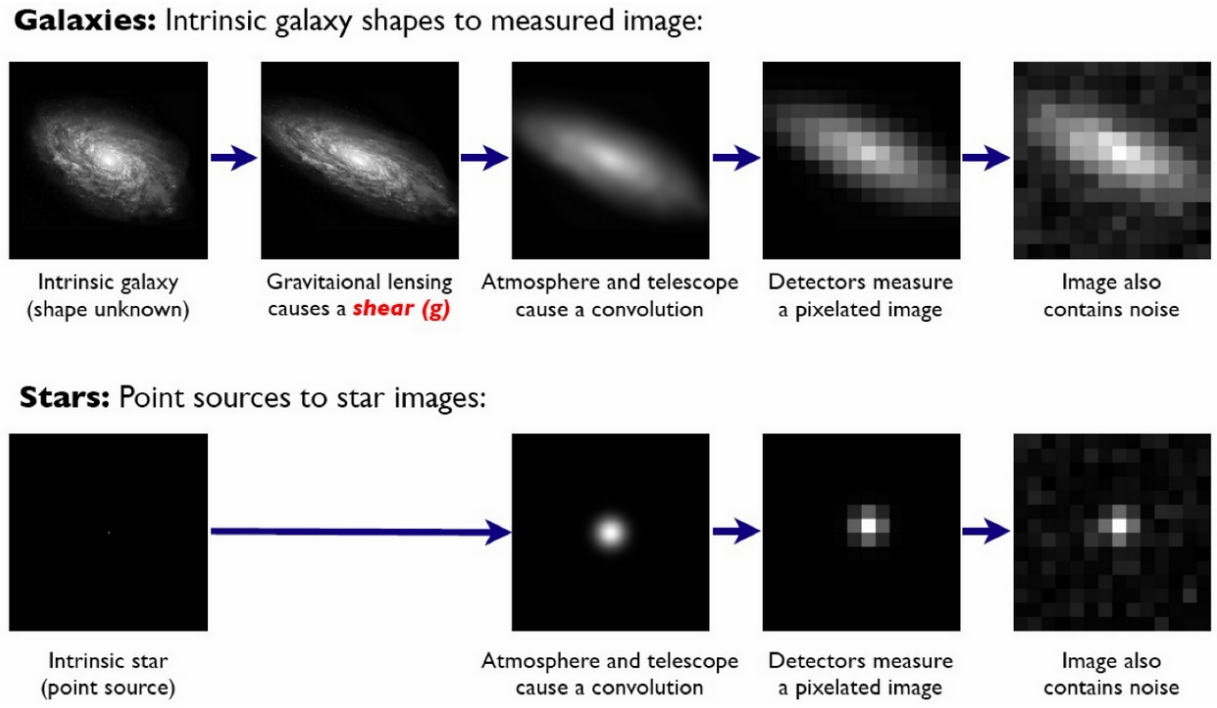

Figure 1.9: Origin of the measured signal from the physical object to the detector readout for the case of a galaxy (top) and for a star (bottom). Figure adapted from Bridle et al. (2009)

Stars, which are point sources, are transferred through the same measurement process, and their apparent shapes and sizes can be used to model the PSF and the characteristics of the detector. This scenario is illustrated on the bottom row of Figure 1.9, where it is explicitly noted that stars do not carry any lensing signal. The two key inputs needed for estimating the shape of a galaxy are the postage stamp images of a galaxy, and the PSF models estimated at the location of the galaxy for each exposure. This latter step is necessary as the PSF varies not only between different observations but also within the focal plane of the telescope. Hence the PSF model is constructed by interpolating the PSFs of a set of tracer stars.

There are two broad approaches for reconstructing the lensed galaxy shape from the above described measurements. The forward modeling approach formulates an analytic model for the intrinsic surface brightness distribution of each galaxy. This model is then transformed to a prediction for each single exposure postage stamp by convolving it with the corresponding PSF, and the result is fitted to the observed galaxy images. In the case of the DES Y1 dataset, the IM3SHAPE algorithm (Zuntz et al., 2013) falls into this category, which assumed a bulge or disk (i.e. De Vaucouleours or exponential) light profile model for the intrinsic morphology of galaxies. The shape estimate of the galaxy is then derived from the ellipticity parameter of the intrinsic light distribution model. In turn, the inverse modeling approach measures shape estimates of the observed smeared and noisy images, and aims to correct these biased estimates for the observational effects. Historically the KSB algorithm (Kaiser et al., 1995) follows this approach. The METACALIBRATION algorithm (Huff & Mandelbaum, 2017; Sheldon & Huff,

2017) in its implementation for DES Y1 also technically falls into this latter category, as it derives a naive ellipticity estimate by fitting a multivariate Gaussian light profile directly to each observed image and then estimates a correction or *response* term for each based on the observational data. However, the theoretical approach of METACALIBRATION can be formulated for both a forward and an inverse modeling way. This method is further introduced in section 1.4.2.4.

1.4.2.3 Systematics and Shear Bias Parametrization

In a survey like the Dark Energy Survey most source galaxies are observed at a rather low signal to noise, which leaves considerable ambiguity about their intrinsic properties: The noisy and pixelated image shown on the rightmost top panel of Figure 1.9 is not at all an unreasonable representation of how a typical source galaxy with complex intrinsic morphology is observed in a survey. In a scenario like this, different a priori assumptions and analysis choices need to be invoked to calculate the ellipticity. Realistically it is inevitable that at least some of our a priori assumptions and choices are incorrect about the galaxies and the observational effects we encounter, and the result of this is a potential bias in the estimated quantities.

In fact the quantity of importance is not the bias in the galaxy shapes themselves, but the potential bias in the gravitational shear $\mathbf{g}$ derived from their ensemble mean ellipticities. To underline the distinction it is illustrative to consider that galaxies have complicated morphologies and their isophotal shapes and major axis directions can change with radius, hence there is no unique way to assign a single ellipticity to a real galaxy.

The traditional way to express shear bias is by writing it as an expansion around the true shear signal (Heymans et al., 2006)

$$g_i^{\mathrm{est}} = (1 + m_i)\, g_i^{\mathrm{true}} + \alpha e_i^{\mathrm{PSF}} + c_i\,, \tag{1.30}$$

where the subscript indicates the i-th component of the complex quantity. g_i^{true} is the true net gravitational shear imprinted on an ensemble of source galaxy images, while g_i^{est} is the estimated one. The parameters m_i and c_i is known as the multiplicative and additive bias respectively, e_i^{PSF} is the ellipticity of the PSF while α is a PSF leakage parameter.

The main types of biases which lead to non-zero m_i, c_i and α_i terms are commonly described in the example of a forward modeling method, however they are equally important albeit less obvious to delineate in inverse modeling approaches (Jarvis et al., 2016; Zuntz et al., 2018; Samuroff et al., 2018). These types are

- *Model bias* which arises when an incorrect light profile model is used to describe the shape of a galaxy (Voigt & Bridle, 2010; Kacprzak et al., 2014).
- *Noise bias* or estimator bias which arises when during the light model fit a single maximum likelihood or maximum posterior point is used instead of the potentially asymmetric likelihood or posterior distribution (Kaiser et al., 2000; Bernstein & Jarvis, 2002; Refregier et al., 2012).
- *Selection bias* or representation bias which relates to preferentially selecting objects whose shapes are aligned in a particular way. For example, selecting for rounder sources can

introduce a bias by promoting the leakage of faint stars (which carry no shear) into the source galaxy catalog, while galaxies whose shapes align with the PSF can be detected with higher significance (Hirata & Seljak, 2003; Miller et al., 2007; Bernstein & Armstrong, 2014a; Fenech Conti et al., 2017).

- *Neighbor bias* which arises from the blending of multiple light sources in a crowded field which can influence the derived shape estimates or the source selection. Furthermore, in the case of unresolved blends, the derived net ellipticity can no longer be directly related the gravitational shear since in reality each component is sheared separately (Hoekstra et al., 2017; Samuroff et al., 2018; Euclid Collaboration et al., 2019).

Given the multitude of systematic effects, the *calibration* of shear estimates is one of the most important tasks in a weak lensing analysis. This can take two conceptual forms:

A Explicitly measuring the m_i, c_i and α_i bias terms of the shear estimation algorithm in an *image simulation*, then using the measured bias values to directly correct the shear in a real observation. This scenario in DES is followed by the IM3SHAPE algorithm.

B Self-calibrating from observational data. This approach in DES is followed by the METACALIBRATION algorithm.

Neither of these approaches is intrinsically correct, and their performance must be *validated*. Given that in the real world there is no suitable reference dataset to be used for the validation of shape estimation methods, validation is also performed using image simulations. There is however a fundamental difference between validating a method in a simulated setting, and calibrating it in the same kind of simulation: The aim of validation is to convince us of the efficacy of a method and estimate its performance and systematic uncertainty. However, calibration directly transfers information from the simulation into the real measurement, therefore there the measured shears are correct only to the extent the simulation was representative of reality.

1.4.2.4 Metacalibration

Based on its excellent performance in a validation setting and the fact that it does not rely on calibration simulations several DES analyses opted to adopt the METACALIBRATION algorithm as their fiducial shear estimate (Sheldon & Huff, 2017; Zuntz et al., 2018). Given its relevance to the present thesis we briefly overview the key features of this algorithm.

The primary input data for METACALIBRATION are the multi-epoch data structures of individual source galaxies described in section 1.4.2.1, which contain both the single exposure images of the objects and also the corresponding PSFs. To increase signal-to-noise, multiple observational bands are processed jointly. From this, first a naive shape estimate is calculated by fitting a multivariate Gaussian model to the observed images using the `ngmix` algorithm. This is naturally a very biased estimate of the shape of the galaxy. However the essence of METACALIBRATION is to provide an innovative way to correct this biased estimator into a very accurate estimator of the mean gravitational shear $\langle \boldsymbol{\gamma} \rangle \approx \langle \mathbf{g} \rangle$. In the following we denote the *estimated*

ellipticity as $\mathbf{e}$ following Huff & Mandelbaum (2017) and Sheldon & Huff (2017), however caution the reader to not interpret this as an intrinsic property of a galaxy, rather as an intermediate quantity used in estimating mean shear.

The essence of METACALIBRATION is to express the estimated shape of a galaxy in the presence of gravitational shear by the Taylor expansion

$$\mathbf{e} = \mathbf{e}|_{\gamma=0} + \frac{\partial \mathbf{e}}{\partial \boldsymbol{\gamma}}\,\boldsymbol{\gamma} + \dots . \tag{1.31}$$

Assuming weak shears and no intrinsic galaxy alignments the estimated mean galaxy shape is then related to the gravitational shear as

$$\langle \mathbf{e} \rangle = \langle \mathsf{R}_\gamma\, \boldsymbol{\gamma} \rangle\,, \tag{1.32}$$

where

$$\mathsf{R}_{\gamma;\,i,j} = \partial e_i \,/\, \partial \gamma_j \tag{1.33}$$

is the *shear response* matrix. Note that in the case of an unbiased shape estimator, R_γ is the unit matrix. From the above, the mean gravitational shear can be then approximated as

$$\langle \boldsymbol{\gamma} \rangle \approx \langle \mathsf{R}_\gamma \rangle^{-1} \langle \mathbf{e} \rangle \approx \langle \mathsf{R}_\gamma \rangle^{-1} \langle \mathsf{R}_\gamma \boldsymbol{\gamma} \rangle\,. \tag{1.34}$$

The essence of METACALIBRATION is to estimate the response term from the observation itself and use it to correct for the bias of the naive shape estimates (Huff & Mandelbaum, 2017; Sheldon & Huff, 2017). In practice this is performed by numerical image manipulations where an observation is first deconvolved from the PSF, a small artificial shear $\gamma_\Delta \approx 0.01$ is applied, and then its is re-convolved with a representation of the PSF to produce a slightly sheared version of the original observation. From these artificially sheared images the shear response is expressed via finite differences in analogy to Equation 1.33 as

$$\langle \mathsf{R}_{\gamma;\,i,j} \rangle = \frac{\langle e_i^+ \rangle - \langle e_i^- \rangle}{\Delta \gamma_j}\,, \tag{1.35}$$

where $\langle e_i^{+\,-} \rangle$ represents the ensemble average of the i-th component of galaxy shapes derived from images with positive or negative artificial shear applied along the j-th shear direction, and $\Delta\gamma = 2\gamma_\Delta$. This shear response term corresponds to the calibrations addressing the model bias and the noise bias of the naive shape estimator, and thus can be used to derive precise and accurate estimates of the mean gravitational shear.

The correction for a different systematic effect: selection bias can also be calculated from the artificially sheared images used in the above step. For this we make use of the fact that selection bias relates to objects preferentially entering or exiting the source catalog depending on their ellipticities. Hence by performing source selection not on the original observed images, but on the artificially sheared ones one can quantify the ellipticity dependence of the selection function. Sheldon & Huff (2017) formulates this in a similar way to Equation 1.35, where the *selection response* is expressed as

$$\langle \mathsf{R}_{S;\,i,j} \rangle = \frac{\langle e_i \rangle^{S+} - \langle e_i \rangle^{S-}}{\Delta \gamma_j}\,, \tag{1.36}$$

where $\langle e_i \rangle^{S+-}$ represents the ensemble average of the i-th component of galaxy shapes derived from the unsheared images but selected based on the sheared images with positive or negative artificial shear applied along the j-th shear direction.

Combining the above response terms the mean shear experienced by a galaxy sample can be expressed as

$$\langle \boldsymbol{\gamma} \rangle \approx \langle \mathsf{R} \rangle^{-1} \langle \mathbf{e} \rangle \,, \tag{1.37}$$

using a combined response term defined as

$$\langle \mathsf{R} \rangle \equiv \langle \mathsf{R}_\gamma \rangle + \langle \mathsf{R}_S \rangle \,. \tag{1.38}$$

The end product of the METACALIBRATION shear estimation algorithm is a catalog containing the naive shape estimates for source galaxies e_i, the four sheared shape estimates of each galaxy e_i^{+-}, and the selection flags derived from the four sheared images of each galaxy. From these data products the mean shapes and the shear and selection response terms can be calculated for the appropriate galaxy ensembles, and hence we can derive an estimate on the mean gravitational shear.

1.4.3 Photometric Redshift Estimation

In addition to source galaxy shapes, the second essential ingredient of weak lensing measurements is an estimate of the lens geometry, for which the distances of source galaxies must be estimated from observational data. In practice this takes the form of estimating the redshifts of source galaxies, which can then be turned into an angular diameter distance through a cosmological model. In an ideal setting, the redshift of a galaxy can be directly estimated from the observed positions of emission or absorption lines in their spectral energy distribution (SED). This type of redshift estimates are commonly referred to as *spectroscopic redshifts*. Measuring the SEDs with sufficient resolution however requires targeted observations and a relatively large investment of observational time for each galaxy, which is not feasible for wide area sky surveys such as DES. Instead, the redshifts of galaxies must be estimated from the integrated flux measured in wide bandpass filters, and this type of estimate is usually denoted as a *photometric redshift* or *photo-z* (Bonnett et al., 2016; Hoyle et al., 2018).

An illustration on the schematic of photometric redshift estimation is shown on Figure 1.10, adapted from Buchs et al. (2019). The figure illustrates the location and spectral throughput of $u, g, r, i, z, Y, J, H, Ks$ observational bandpass filters used in the deep fields of DES. Of these only the g, r, i, z filters are fully measured in the wide field area of DES. The top panel of the Figure 1.10 illustrates the imprint of redshift on the fluxes within the different filters with a set of example SEDs overlayed corresponding to a quenched elliptical galaxy at different redshifts. The ratio of fluxes in the different bandpass filters corresponds to the colors of the galaxy, and the behavior shown in this panel is loosely referred to as the color-redshift relation. This relation is the basis of most widely used photometric redshift algorithms[11], including the the ones employed

[11] While there are photo-z methods which employ explicit galaxy SED templates, and methods which attempt to learn the color-redshift mapping in a non-parameteric way, the fundamental properties and limitations of these methods maare very similar.

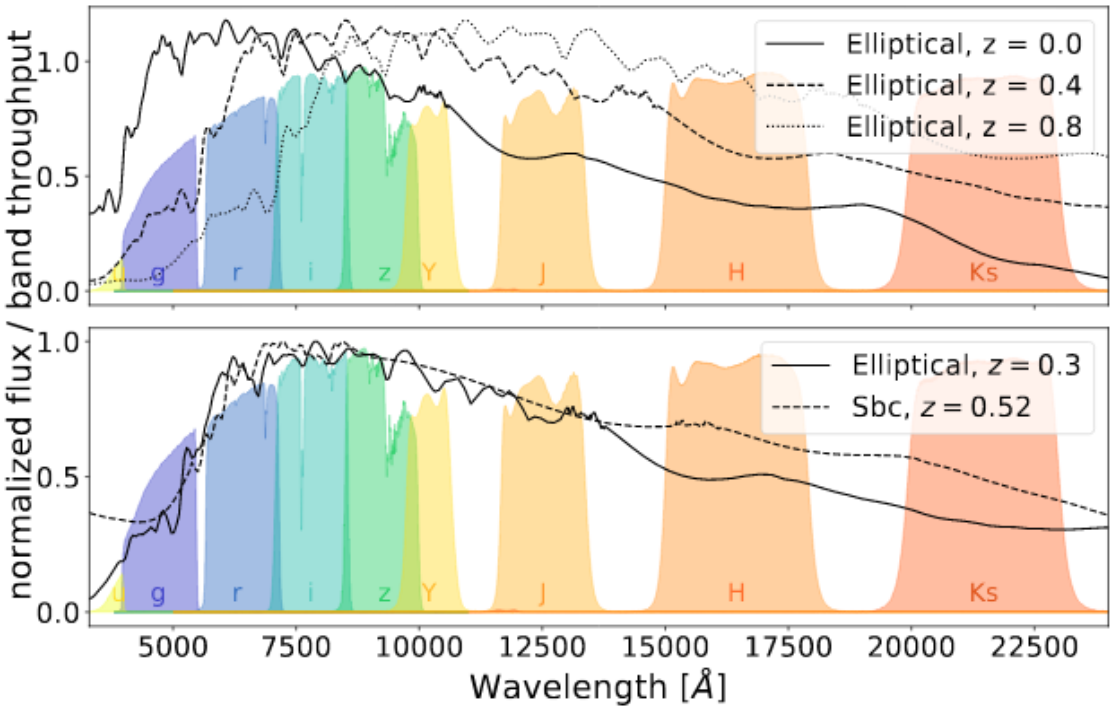

Figure 1.10: Redshift estimation schematic for a wide band sky survey. *Top:* Galaxy SEDs at different redshifts can be distinguished based on their relative fluxes (colors) measured in different bandpass filters, also known as the color-redshift relation. *Bottom:* Illustration of a degeneracy in the color-redshift relation where different galaxy types at different redshifts can produce similar fluxes in a bandpass filter. Figure adapted from Buchs et al. (2019).

in the later chapters of this thesis (i.e. `BPZ` Benítez, 2000, `EAzY` Brammer et al., 2008), therefore we restrict our discussion to the fundamental properties of these methods.

Observational uncertainties in the flux (color) measurement manifest as an uncertainty in the galaxy redshifts derived from the color-redshift relation. For this reason it is beneficial to instead of considering point estimates, that is single redshift values, introduce the redshift probability distribution function $p(z)$ of a galaxy whose shape and width captures said uncertainties. This formalism allows us to incorporate a further complication related to photometric redshift estimates, namely that the color-redshift relation is degenerate between galaxy type (SED) and redshift. This fact is illustrated on the bottom panel of Figure 1.10, where two distinct galaxy types located at different redshifts possess a very similiar observed SED. For this example the degeneracy is only broken by the infrared bands, which means that these two galaxy types and redshifts would not be distinguishable in the DES wide field survey, and there a bimodal $p(z)$ would have to be assigned. In fact due to the wide variety of intrinsic galaxy SEDs and the considerable redshift range probed by current sky surveys (with source galaxies up to $\langle z \rangle \approx 1.1$) these degeneracies are not localized to specific examples. Hence when estimating the photometric redshifts of a galaxy the redshift estimate must be marginalized over all galaxy types and redshfits, and the resulting photo-z $p(z)$ usually spans a relatively large redshift range (Hoyle et al., 2018).

Where does the information come from to correctly perform the above marginalization? In broad terms DES and other similar sky surveys use spectroscopic or narrow-band reference or training datasets such as the COSMOS field (Scoville, 2003) or the DES deep fields where the mapping between colors, magnitudes and redshift can be learned. Learning this mapping from observational training data, and then applying it to estimate the $p(z)$ of galaxies in the full survey dataset is the core task of photometric redshift estimation. Within this topic the bulk of the effort is usually spent on validating the redshift estimates, which in brief relates to quantifying their systematic uncertainties and correcting their biases based on tests against other alternative redshift estimates (Bonnett et al., 2016; Davis et al., 2018; Hoyle et al., 2018; Buchs et al., 2019).

In the context of data products, the photometric redshift estimates are derived based on the colors and magnitudes of the galaxies identified in DES. The base photometric catalog for this is by default the DES GOLD catalog, however for weak lensing measurements in some cases the photometry catalog constructed by the METACALIBRATION algorithm is also used. The contents of a photometric redshift catalog in DES include for each galaxy the mean redshift estimate $\langle z \rangle$ and Monte-Carlo samples drawn form the $p(z)$.

1.5 A Primer on Cluster Cosmology

As already introduced in section 1.1.3, the mass function of halos is sensitive to the parameters and properties of the cosmological model. In an empirical setting however massive halos are not observed directly, instead we see them as groups and clusters of galaxies. This means that theoretical or numerical predictions on the halo mass function must be related to the distribution of clusters as function of directly observed or measured quantities. In the following I briefly outline the analysis strategy adopted by the Dark Energy Survey Cluster Cosmology working group (DES Collaboration et al., 2020) in its year 1 analysis.

In the optical dataset of DES Y1 galaxy clusters are identified via the redMaPPer algorithm (section 1.4.1) and characterized by the observed optical richness parameter $\lambda^{\rm obs}$, whereas the redshifts $z^{\rm obs}$ of clusters are estimated from the ensemble of cluster galaxies. The superscript obs is explicitly written out in this section to indicate that both of these quantities are considered to be a noisy and potentially biased measurement compared to the intrinsic true values of the clusters $\lambda^{\rm true}$ and $z^{\rm true}$, although in the case of cluster redshifts this is estimated to be a minor effect. The connection between observed cluster richness and mass is not a-priory known to sufficient precision to enable the direct use of the richness function of galaxy clusters for cosmological inference. The masses of galaxy clusters must be calibrated externally, in the case of DES this can be performed via weak lensing measurements.

A characteristic property of weak lensing measurements is that the signal-to-noise is low, therefore either an exceptionally massive set of clusters or a larger statistical ensemble of lower mass clusters is needed to reach precise and accurate mass constraints. DES provides very large catalogs of clusters within the mass range from $M_{200m} \approx 10^{14}\mathrm{M}_\odot$ to $\mathrm{M}_{200m} \approx 10^{14.8}\mathrm{M}_\odot$, which are suitable for the ensemble lensing strategy of mass calibration. This means that the cosmology analysis must be formulated for a binned analysis where the only direct inputs are mean cluster properties within bins of observables, that is in bins of richness and redshift.

In this strategy the two primary measurements adopted by (DES Collaboration et al., 2020) are the mean number of clusters $\langle N\rangle$ in a bin richness and redshift, and the mean weak lensing cluster mass $\langle M\rangle$ of the corresponding cluster sample. These quantities are related to the halo mass function such that

$$\langle N\rangle = \int_0^\infty \mathrm{d}\, z^{\rm true} \int_{z_{\rm min}}^{z_{\rm max}} \mathrm{d}z^{\rm obs} \int_{\lambda_{\rm min}}^{\lambda_{\rm max}} \mathrm{d}\lambda^{\rm obs}\ \langle n \,|\, \lambda^{\rm obs},\, z^{\rm true}\rangle \,\frac{\mathrm{d}V}{\mathrm{d}z^{\rm true}}\, P(z^{\rm obs}\,|\,z^{\rm true}) \tag{1.39}$$

and

$$\langle n \,|\, \lambda^{\rm obs},\, z^{\rm true}\rangle = \int_0^\infty \mathrm{d}M\, \frac{\mathrm{d}n}{\mathrm{d}M}\, P(\lambda^{\rm obs}|M,\, z^{\rm true}) \tag{1.40}$$

where $\mathrm{d}n/\mathrm{d}M$ is our theoretical or numerical estimate of the halo mass function and $P(\lambda^{\rm obs}|M,\, z^{\rm true})$ is a conditional probability of the richness – mass – redshift relation. $P(z^{\rm obs}\,|\,z^{\rm true})$ captures uncertainty in cluster redshift estimates, $\lambda_{\rm min}$, $\lambda_{\rm max}$ and $z_{\rm min}$ and $z_{\rm max}$ refer to the boundaries of the richness and redshift bin respectively, and

$$\frac{\mathrm{d}V}{\mathrm{d}z} = \Omega_{\rm mask}(z) c H^{-1}(z)\, D_c^2(z) \tag{1.41}$$

refers to the comoving volume element of the sky survey with angular footprint $\Omega_{\rm mask}$.

From the same ingredients the mean cluster mass can be expressed as

$$\langle M \rangle = \frac{1}{\langle N \rangle} \int_0^\infty \mathrm{d}\, z^{\rm true} \int_{z_{\rm min}}^{z_{\rm max}} \mathrm{d}z^{\rm obs} \int_{\lambda_{\rm min}}^{\lambda_{\rm max}} \mathrm{d}\lambda^{\rm obs}\ \langle nM \,|\, \lambda^{\rm obs},\, z^{\rm true} \rangle \, \frac{\mathrm{d}V}{\mathrm{d}z^{\rm true}}\, P(z^{\rm obs} \,|\, z^{\rm true}) \quad (1.42)$$

where

$$\langle nM | \lambda^{\rm obs},\, z^{\rm true} \rangle = \int_0^\infty \mathrm{d}M\, \frac{\mathrm{d}n}{\mathrm{d}M}\, M\, P(\lambda^{\rm obs} | M,\, z^{\rm true})\,. \quad (1.43)$$

Using the above formulations, the predictions of cosmological theory can be compared with observations of the abundance and masses of galaxy clusters. The further details of this model and discussing its various analysis choices are beyond the scope of this thesis, and hence we refer the reader to the methods review of Costanzi et al. (2019). The following chapters of original research presented in this thesis relate primarily to the precise and accurate measurement of $\langle M \rangle$.

Chapter 2

Weak Lensing Mass Calibration of redMaPPer Galaxy Clusters

Scientific context The aim of this chapter is to calibrate the masses of galaxy clusters via weak lensing for the clusters identified in the DES year 1 dataset using the redMaPPer algorithm. These measurements are then used in the Dark Energy Survey to derive cosmological constraints from the abundances and masses of galaxy clusters. This work fits into the broader cluster cosmology analysis the following way:

- The redMaPPer cluster finder is run on the DES Y1 data and a catalog is constructed of the identified clusters along with their estimated redshift, and richness – an optical mass proxy (see section 1.4.1). Defining a binning scheme in richness and redshift we can measure the detected number of clusters as function of observables, which is in turn can be related to the mean number of clusters $\langle N \rangle$ (Equation 1.39) expected from the halo mass function, the cosmological model and the survey strategy.

- We develop a cluster weak lensing measurement framework for DES Y1 dataset tailored to the format and performance of the source galaxy catalog and photometric redshift information available in the survey, and measure mean $\Delta\Sigma$ profiles of clusters in bins of richness and redshift. We then define a model for the mass profile and for a series of modulating systematic effects and estimate the mean cluster mass corresponding to each bin of redshift and richness. This mean cluster mass can be then directly related to the to the $\langle M \rangle$ expressed in Equation 1.42 expected from the cosmological model.

- Using the derived mean cluster masses we fit the mean of the richness–mass relation. This is done to quantify the performance of the measurement, by propagating the statistical and systematic uncertainty of the mass estimates across the different bins into an uncertainty estimate on the mean observable–mass scaling, which is the leading systematic of cosmology analyses.

Therefore the first two of the above points constitute the direct inputs of the cosmology analysis performed in DES Collaboration et al. (2020), while the last point relates to estimating the performance of the mass calibration.

Authorship and publication The research presented in this chapter has been conducted within the Dark Energy Survey, and was published as *Dark Energy Survey Year 1 results: weak lensing mass calibration of redMaPPer galaxy clusters* in the Monthly Notices of the Royal Astronomical Society, volume 482, pages 1352–1378 (McClintock, Varga et al., 2019).

This project has been lead by two corresponding authors with equal contribution: T. McClintock and T. N. Varga (myself). According to DES publication policy the additional authors of the paper are divided into a first tier with direct contribution to the research in terms of data analysis or substantive discussion, second tier authors whose work indirectly enabled the analysis and whom are listed alphabetically, and a so-called *builder*-tier also listed alphabetically and which consists of people whose work contributed to DES infrastructure as a whole.

The primary contributions of the corresponding authors is divided as the following:

- T. N. Varga: Data processing and preparation (section 2.2.3, section 2.2.4), weak lensing profile measurement (section 2.3.1), systematic tests (section 2.4.1, section 2.4.1.2, section 2.4.1.3), cluster member contamination estimates (section 2.4.2), empirically measured covariance model components (section 2.3.2.1, section 2.3.2.2).

- T. McClintock: Cluster mass fitting and likelihood (section 2.5), mass-richness scaling fit (section 2.6), analytically modeled covariance model components (section 2.3.2.3, section 2.3.2.4).

Among the first tier authors E. Rykoff created the redMaPPer catalog for the DES Y1 dataset (section 2.2.2), while T. Shin performed the calculation of the multiplicative redshift bias correction (section 2.4.3). We furthermore thank the advisory contribution of D. Gruen, E. Rozo, and S. Seitz. In addition to this P. Melchior and J. P. Dietrich provided useful discussion on modeling cluster triaxiality and projection systematics (section 2.5.4.2), Y. Zhang provided measurements on the centering properties of redMaPPer galaxy clusters from multi-wavelength follow up (Zhang et al., 2019b), J. DeRose contributed the numerical simulations used in section 2.5.4.1, and E. Sheldon contributed useful discussions about the use of metacalibration in the cluster weak lensing scenario. All first tier authors contributed to the discussion section of the paper. In accordance with DES publication policy this paper has gone through internal review and collaboration wide review before the submitting to the journal.

The following part of this chapter reproduces the paper as it was accepted for publication in the journal with minor edits to suit the thesis format, while keeping the substantive part of the work unchanged. The acknowledgments and appendix of the original paper are reproduced in Appendix A.

Abstract We constrain the mass–richness scaling relation of redMaPPer galaxy clusters identified in the Dark Energy Survey Year 1 data using weak gravitational lensing. We split clusters into 4×3 bins of richness λ and redshift z for $\lambda \geq 20$ and $0.2 \leq z \leq 0.65$ and measure the mean masses of these bins using their stacked weak lensing signal. By modeling the scaling relation as $\langle M_{200\mathrm{m}}|\lambda, z\rangle = M_0(\lambda/40)^F((1+z)/1.35)^G$, we constrain the normalization of the scaling relation at the 5.0 per cent level as $M_0 = [3.081 \pm 0.075(\mathrm{stat}) \pm 0.133(\mathrm{sys})] \cdot 10^{14}$ $\mathrm{M}_\odot$ at $\lambda = 40$ and $z = 0.35$. The richness scaling index is constrained to be $F = 1.356 \pm 0.051$ (stat) $\pm$ 0.008 (sys) and the redshift scaling index $G = -0.30 \pm 0.30$ (stat) $\pm$ 0.06 (sys). These are the tightest measurements of the normalization and richness scaling index made to date. We use a semi-analytic covariance matrix to characterize the statistical errors in the recovered weak lensing profiles. Our analysis accounts for the following sources of systematic error: shear and photometric redshift errors, cluster miscentering, cluster member dilution of the source sample, systematic uncertainties in the modeling of the halo–mass correlation function, halo triaxiality, and projection effects. We discuss prospects for reducing this systematic error budget, which dominates the uncertainty on M_0. Our result is in excellent agreement with, but has significantly smaller uncertainties than, previous measurements in the literature, and augurs well for the power of the DES cluster survey as a tool for precision cosmology and upcoming galaxy surveys such as LSST, Euclid and WFIRST.

2.1 Introduction

Galaxy clusters have the potential to be the most powerful cosmological probe (Dodelson et al., 2016). Current constraints are dominated by uncertainties in the calibration of cluster masses (e.g., Mantz et al., 2015; Planck Collaboration et al., 2016; Rozo et al., 2010). Weak lensing allows us to determine the mass of galaxy clusters: gravitational lensing of background galaxies by foreground clusters induces a tangential alignment of the background galaxies around the foreground cluster. This alignment is a clear observational signature predicted from clean, well-understood physics. Moreover, the resulting signal is explicitly sensitive to all of the cluster mass, not just its baryonic component, and is insensitive to the dynamical state of the cluster. For all these reasons, weak lensing is the most robust method currently available for calibrating cluster masses. It is therefore not surprising that the community has invested in a broad range of weak lensing experiments specifically designed to calibrate the masses of galaxy clusters (von der Linden et al., 2014b,a; Applegate et al., 2014a; Hoekstra et al., 2015; Okabe & Smith, 2016; Mantz et al., 2015; Melchior et al., 2017; Simet et al., 2017b; Murata et al., 2017; Dietrich et al., 2019; Miyatake et al., 2018; Medezinski et al., 2018a).

The Dark Energy Survey (DES) is a 5,000 square degree photometric survey of the southern sky. It uses the 4-meter Blanco Telescope and the Dark Energy Camera (Flaugher et al., 2015) located at the Cerro Tololo Inter-American Observatory. As its name suggests, the primary goal of the DES is to probe the physical nature of dark energy, in addition to constraining the properties and distribution of dark matter. Owing to its large area, depth, and image quality, at its conclusion DES will support optical identification of $\sim 100,000$ galaxy clusters and groups up to redshift $z \approx 1$. We use galaxy clusters identified using the redMaPPer algorithm (Rykoff et al.,

2014), which assigns each cluster a photometric redshift and optical richness λ of red galaxies. To fully utilize these clusters, one must understand mass-observable relations (MORs), such as that between cluster mass and optical richness. Weak lensing can establish this relation – with high statistical uncertainty for individual clusters, but low systematic uncertainty in the mean mass scale derived from the joint signal of large samples.

In this work, we use stacked weak lensing to measure the mean galaxy cluster mass of redMaPPer galaxy clusters identified in DES Year 1 (Y1) data. We use these data to calibrate the mass–richness–redshift relation of these clusters. In Melchior et al. (2017) we provided a first calibration of this relation using DES Science Verification (SV) data. There, we were able to achieve a 9.2 per cent statistical and 5.1 per cent systematic uncertainty. Here, we update that result using the first year of regular DES observations, incorporating a variety of improvements to the analysis pipeline. Our results provide the tightest, most accurate calibration of the richness–mass relation of galaxy clusters to date, at 2.4 per cent statistical and 4.3 per cent systematic uncertainty.

The structure of this section is as follows. In section 2.2, we introduce the DES Y1 data used in this work. In section 2.3 we describe our methodology for obtaining ensemble cluster density profiles from stacked weak lensing shear measurements, with a focus on updates relative to Melchior et al. (2017). A comprehensive set of tests and corrections for systematic effects is presented in section 2.4. The model of the lensing data and the inferred stacked cluster masses are given in section 2.5. The main result, the mass–richness–redshift relation of redMaPPer clusters in DES, is presented in section 2.6. We compare our results to other published works in the literature in section 2.7, discuss systematic improvements made in this work compared to Melchior et al. (2017) in section 2.8, and conclude in section 3.5. In Appendix A.2 we present the DES Y1 redMaPPer catalog used in this work for public use. Supplementary information on the analysis is given in additional appendices.

Unless otherwise stated, we assume a flat ΛCDM cosmology with $\Omega_\mathrm{m} = 0.3$ and $H_0 = 70$ km s^{-1} Mpc^{-1}, with distances defined in physical coordinates, rather than comoving. Finally, unless otherwise noted all cluster masses refer to $M_{200\mathrm{m}}$. That is, cluster mass is defined as the mass enclosed within a sphere whose average density is 200 times higher than the mean cosmic matter density $\bar{\rho}_m$ at the cluster's redshift, matching the mass definition used in the cosmological analyses that make use of our calibration.

2.2 The DES year 1 data

DES started its main survey operations in 2013, with the Year One (Y1) observational season running from August 31, 2013 to February 9, 2014 (Drlica-Wagner et al., 2018). During this period 1839 deg^2 of the southern sky were observed in three to four tilings in each of the four DES bands g, r, i, z, as well as $\sim$1800 deg^2 in the Y-band. The resulting imaging is shallower than the SV data release but covers a significantly larger area. In this study we utilize approximately 1500 deg^2 of the main survey, split into two large non-contiguous areas. This is a reduction from the 1800 deg^2 area due to a series of veto masks. These masks include masks for bright stars and the Large Magellanic Cloud, among others. The two non-contiguous areas are the "SPT"

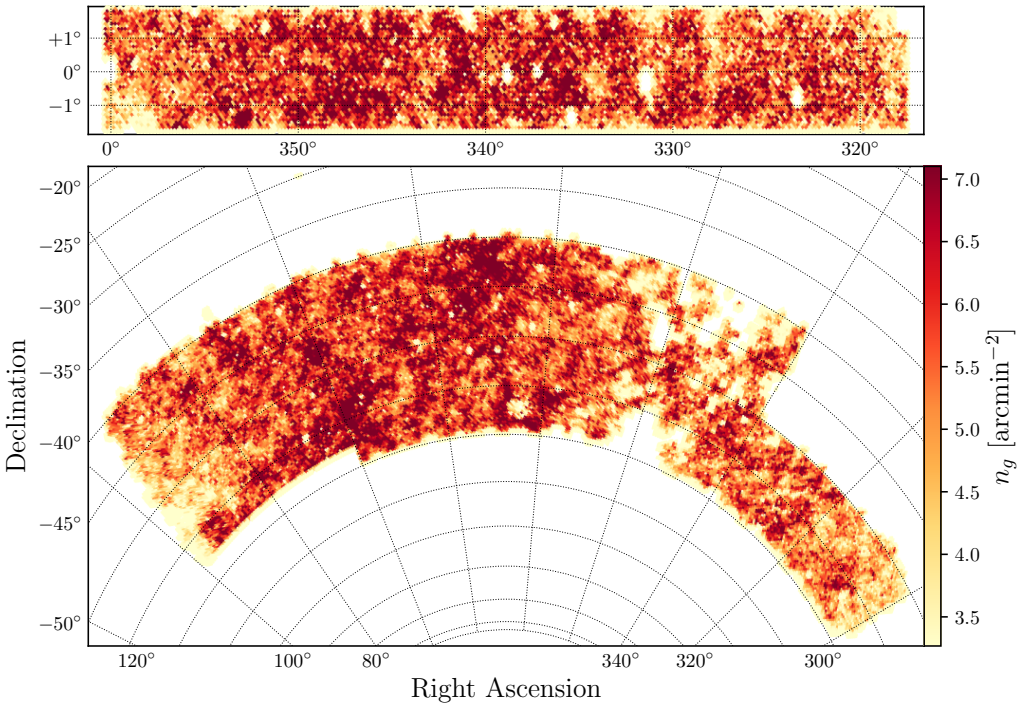

Figure 2.1: Surface density of source galaxies in the METACALIBRATION catalog within the DES Y1 footprint in the "S82" field (*top*) and the "SPT" field (*bottom*).

area (1321 deg^2), which overlaps the footprint of the South Pole Telescope Sunyaev-Zel'dovich Survey (Carlstrom et al., 2011), and the "S82" area (116 deg^2), which overlaps the Stripe-82 deep field of the Sloan Digital Sky Survey (SDSS; Annis et al., 2014). The DES Y1 footprint is shown in Figure 2.1.

In the following we briefly describe the main data products used in this analysis, and refer the reader to the corresponding papers for more details. The input photometric catalog, as well as the photometric redshift and weak lensing shape catalogs used in this study have already been employed in the cosmological analysis combining galaxy clustering and weak lensing by the DES collaboration (Abbott et al., 2018).

2.2.1 Photometric Catalog

Input photometry for the redMaPPer cluster finder (section 2.2.2) and photometric redshifts (section 2.2.4) was derived from the DES Y1A1 Gold catalog (Drlica-Wagner et al., 2018). Y1A1 Gold is the science-quality internal photometric catalog of DES created to enable cosmological analyses. This data set includes a catalog of objects as well as maps of survey depth and foreground masks, and star-galaxy classification. In this work we make use of the multi-epoch, multi-object fitting (MOF) composite model (`CM`) galaxy photometry. The MOF photometry simultaneously fits a psf-convolved galaxy model to all available epochs and bands for each object, while subtracting and masking neighbors. The typical 10σ limiting magnitude inside $2''$ diameter

apertures for galaxies in Y1A1 Gold using MOF `CM` photometry is $g \approx 23.7$, $r \approx 23.5$, $i \approx 22.9$, and $z \approx 22.2$. Due to its low depth and calibration uncertainty, we do not use Y band photometry for shape measurement or photometric redshift estimation.

The galaxy catalog used for the redMaPPer cluster finder is constructed as follows. Bad objects that are determined to be catalog artifacts, including having unphysical colors, astrometric discrepancies, and PSF model failures are rejected (Section 7.4 Drlica-Wagner et al., 2018). Galaxies are then selected via the more complete `MODEST_CLASS` classifier (Section 8.1 Drlica-Wagner et al., 2018). Only galaxies that are brighter in z band than the local $10\,\sigma$ limiting magnitude are used by redMaPPer. The average survey limiting magnitude is deep enough to image a $0.2\,L^*$ galaxy at $z \approx 0.7$. Finally, we remove galaxies in regions that are contaminated by bright stars, bright nearby galaxies, globular clusters, and the Large Magellanic Cloud.

2.2.2 Cluster catalog

We use a volume limited sample of galaxy clusters detected in the DES Y1 photometric data using the redMaPPer cluster finding algorithm v6.4.17 (Rykoff et al., 2014, 2016). This redMaPPer version is identical to the v6.3 algorithm described in Rykoff et al. (2016) but updated to accommodate the new redMaGiC galaxy catalogs (Rozo et al., 2016; Elvin-Poole et al., 2018). The redMaGiC algorithm selects luminous red galaxies in such a way as to minimize the photometric redshift uncertainties, which serves as an input to redMaPPer in order to construct a red sequence (Rozo et al., 2016).

Two versions of the redMaPPer cluster catalog are generated: a "flux limited" version, which includes high redshift clusters for which the richness requires extrapolation along the cluster luminosity function, and one that is locally volume-limited. By "locally volume-limited" we mean that at each point in the sky, a galaxy cluster is included in the sample if and only if all cluster galaxies brighter than the luminosity threshold used to define cluster richness in redMaPPer lies above $10\,\sigma$ in z, $5\,\sigma$ in i and r, and $3\,\sigma$ in g according to the survey MOF depth maps (Drlica-Wagner et al., 2018). That is, no extrapolation in luminosity is required when estimating cluster richness. At the threshold the galaxy sample is $> 90 - 95$ per cent complete. It is this volume-limited cluster sample that is used in follow-up work deriving cosmological constraints from the abundance of galaxy clusters. Consequently, we focus exclusively on this volume-limited sample in this work. It contains more than 76,000 clusters down to $\lambda > 5$, of which more than 6,500 are above $\lambda = 20$. The format of the catalogs are described in section A.2.

redMaPPer identifies galaxy clusters as overdensities of red-sequence galaxies. Starting from an initial set of spectroscopic seed galaxies, the algorithm iteratively fits a model for the local red-sequence, and finds cluster candidates while assigning a membership probability to each potential member. Clusters are centered on bright galaxies selected using an iteratively self-trained matched-filter method. The method allows for the inherent ambiguity of selecting a central galaxy by assigning a probability to each galaxy of being the central galaxy of the cluster. The final membership probabilities of all galaxies in the field are assigned based on spatial, color, and magnitude filters.

The distribution of cluster richness and redshift of the DES volume-limited cluster sample is shown in Figure 2.2. The richness estimate λ is the sum over the membership probabilities of all

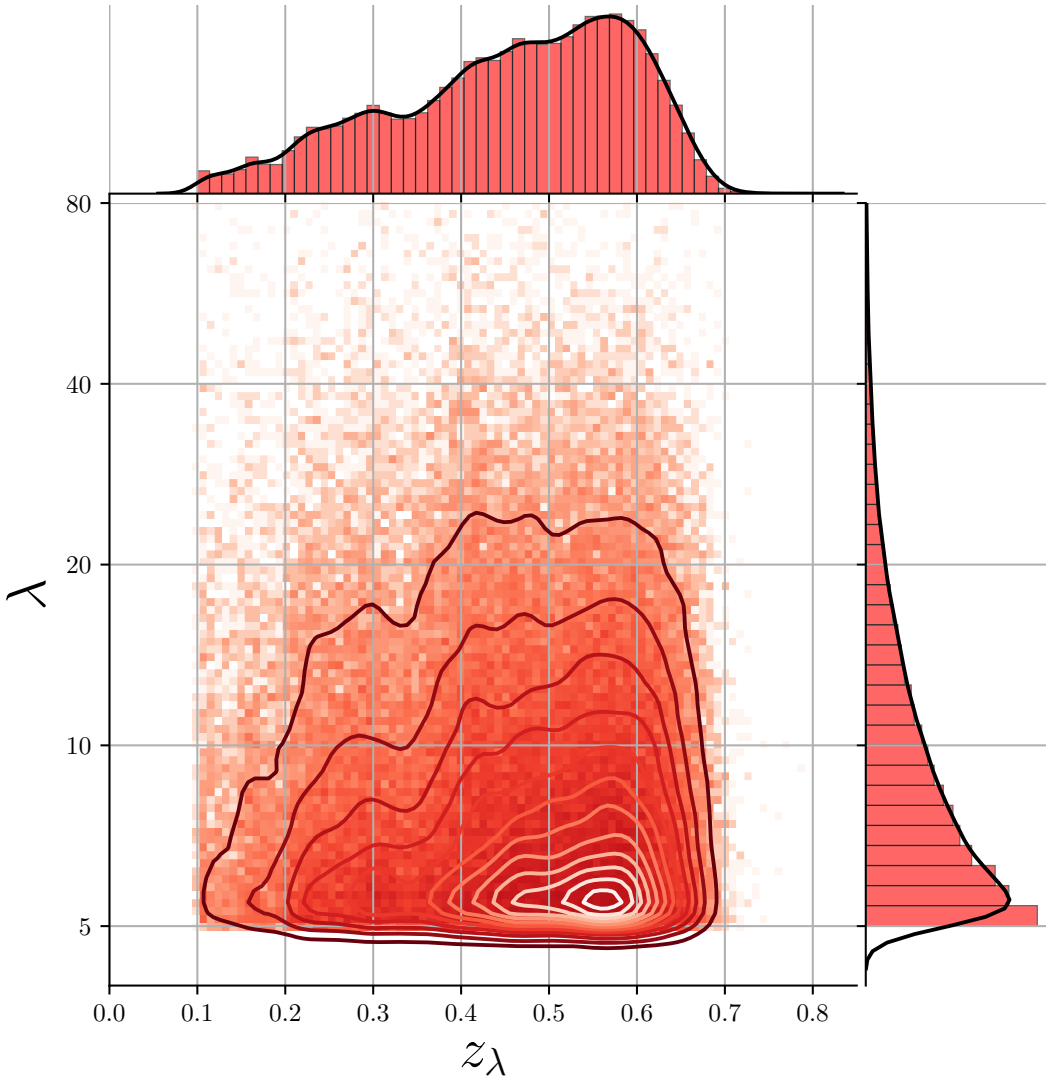

Figure 2.2: Redshift–richness distribution of redMaPPer clusters in the volume limited DES Y1 cluster catalog, overlaid with density contours to highlight the densest regions. At the top and on the right are histograms of the projected quantities, z_λ and λ, respectively, with smooth kernel density estimates overlaid.

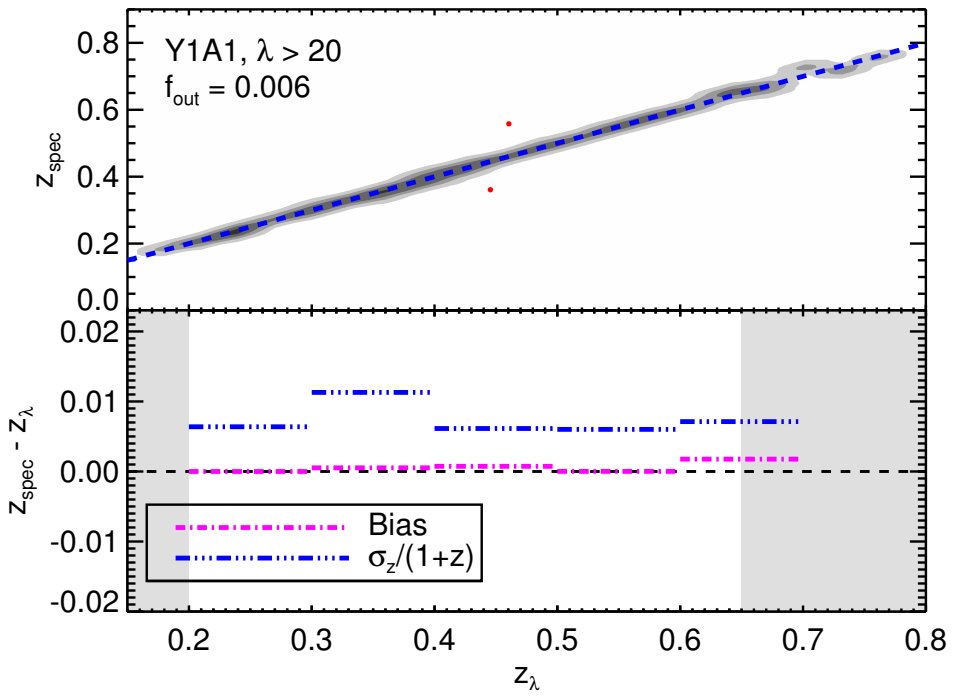

Figure 2.3: Photometric redshift performance of the DES Y1 redMaPPer cluster catalog, as evaluated using available spectroscopy (333 clusters). *Upper panel*: Gray contours are 3σ confidence intervals, and the two red dots are the only 4σ outliers, caused by miscentering on a foreground/background galaxy. *Lower Panel*: photo-z bias and uncertainty. The comparatively large uncertainty from $0.3 < z < 0.4$ is due to a filter transition.

galaxies within a pre-defined, richness–dependent projected radius R_λ. The radius R_λ is related to the cluster richness via $R_\lambda = 1.0(\lambda/100)^{0.2}\ h^{-1}$Mpc. This relation was found to minimize the scatter between richness and X-ray luminosity in Rykoff et al. (2012). A redshift estimate for each cluster is obtained by maximizing the probability that the observed color-distribution of likely members matches the self-calibrated red-sequence model of redMaPPer.

Figure 2.3 shows the photometric redshift performance of the DES Y1 volume-limited redMaPPer cluster sample. The photometric redshift bias and scatter are calculated by comparing the photometric redshift of the clusters to the spectroscopic redshift of the central galaxy of the cluster, where available. Unfortunately, the small overlap with existing spectroscopic surveys means that our results are limited by small-number statistics: there are only 333 galaxy clusters with a spectroscopic central galaxy, and only 34 (six) with redshift $z \geq 0.6$ ($z \geq 0.65$). Nevertheless, the photometric redshift performance is consistent with our expectations: our redshifts are very nearly unbiased, and have a remarkably tight scatter — the median value of $\sigma_z/(1+z)$ is ≈ 0.006. An upper limit for the photometric redshift bias of 0.003 is consistent with our data.

Of particular importance to this work is the distribution of miscentered clusters – both the frequency and severity of their miscentering. Based on the redMaPPer centering probabilities,

we would expect $\approx$ 80 per cent of the clusters to be correctly centered, meaning the most likely redMaPPer central galaxy is at the center of the potential well of the host halo. In practice, the fraction of correctly centered galaxy clusters is closer to $\approx$ 70 per cent, as estimated from a detailed comparison of the redMaPPer photometric centers to the X-ray centers of redMaPPer clusters for which high-resolution X-ray data is available (Zhang et al., 2019b; von der Linden et al., 2018). The expected impact of this miscentering effect, and the detailed model for the miscentered distribution from Zhang et al. (2019b); von der Linden et al. (2018) is described in section 2.5.2.

2.2.3 Shear catalogs

Our work uses the DES Y1 weak lensing galaxy shape catalogs presented in Zuntz et al. (2018). Two independent catalogs were created: METACALIBRATION (Sheldon & Huff, 2017; Huff & Mandelbaum, 2017) based on NGMIX (Sheldon, 2015), and IM3SHAPE (Zuntz et al., 2013). Both pass a multitude of tests for systematics, making them suitable for cosmological analyses. While the Y1 data is shallower than the DES SV data, improvements in the shear estimation pipelines and overall data quality enabled us to reach a number density of sources similar to that from DES SV data (Jarvis et al., 2016).

In this study we will focus exclusively on the METACALIBRATION shear catalog because of its larger effective source density (6.28 arcmin^{-2}) compared to the IM3SHAPE catalog (3.71 arcmin^{-2}). The difference mainly arises because METACALIBRATION utilizes images taken in r, i, z bands, whereas IM3SHAPE relies exclusively on r-band data. In the METACALIBRATION shear catalog the fiducial shear estimates are obtained from a single Gaussian fit via the NGMIX algorithm. As a supplementary data product METACALIBRATION provides (g, r, i, z)-band fluxes and the corresponding error estimates for objects using its internal model of the galaxies.

Galaxy shape estimators, such as the NGMIX model-fitting procedure used for METACALIBRATION, are subject to various sources of systematic errors. For a stacked shear analysis, the dominant problem is a multiplicative bias, i.e. an over- or underestimation of gravitational shear as inferred from the mean tangential ellipticity of lensed galaxies. This bias needs to be characterized and corrected. Traditionally, this is done using simulated galaxy images – with the critical limitation that simulations never fully resemble the observations.

The METACALIBRATION catalog, in contrast, uses the galaxy images themselves to de-bias shear estimates. Specifically, each galaxy image is deconvolved from the estimated point spread function (PSF), and a small positive and negative shear is applied to the deconvolved image in both the $\hat{e}_1$ and $\hat{e}_2$ directions. The resulting images are then convolved once again with a representation of the PSF, and an ellipticity is estimated for these new images (Zuntz et al., 2018). These new measurements can be used to directly estimate the response of the ellipticity measurement to a gravitational shear using finite difference derivatives:

$$\mathsf{R}_\gamma = \frac{\partial \mathbf{e}}{\partial \gamma} \,. \tag{2.1}$$

Selection effects can also be accounted for by examining the response of the selections to shear. The application of a weight when calculating the mean shear over an ensemble is effectively a

type of smooth selection, and is accounted for in the same way. We describe this effect with a selection response $\mathsf{R}_{\rm sel}$, which leads to the response-corrected mean shear estimate

$$\langle\gamma\rangle \approx \langle\mathsf{R}\rangle^{-1}\langle\mathsf{R}\cdot\gamma\rangle \approx \langle\mathsf{R}\rangle^{-1}\langle\mathbf{e}\rangle \tag{2.2}$$

from biased measurements $\mathbf{e}$ with a joint response $\mathsf{R} \approx \mathsf{R}_\gamma + \mathsf{R}_{\rm sel}$ (Sheldon & Huff, 2017).

R is a 2×2 Jacobian matrix for the two ellipticity components e_1, e_2 in a celestial coordinate system. For the METACALIBRATION mean shear measurements in this work, we calculate the response of mean *tangential* shear on mean tangential ellipticity. R is close to isotropic on average, which is why other recent weak lensing analyses (Troxel et al., 2017; Prat et al., 2018; Gruen et al., 2018; Chang et al., 2018a) have assumed it to be a scalar. For the larger tangential shears measured on small scales around clusters, however, we account for the fact that the response might not be quite isotropic by explicitly rotating it to the tangential frame.

Tangential ellipticity $e_{\rm T}$ is related to e_1, e_2 (and likewise γ_T to γ_1 and γ_2) by

$$e_{\rm T} = -e_1\cos(2\phi) - e_2\sin(2\phi)\,, \tag{2.3}$$

where ϕ is the polar angle of the source in a coordinate system centered on the lens. For the shear response, the corresponding rotation is derived from Equation 2.1 and Equation 2.3 as

$$\begin{aligned}\mathsf{R}_{\gamma,{\rm T}} = \mathsf{R}_{\gamma,11}\cos^2(2\phi) + \mathsf{R}_{\gamma,22}\sin^2(2\phi) + \\ \left(\mathsf{R}_{\gamma,12} + \mathsf{R}_{\gamma,21}\right)\sin(2\phi)\cos(2\phi)\,.\end{aligned} \tag{2.4}$$

For the METACALIBRATION selection response, no such rotation can be performed as the term itself is only meaningful for ensembles of galaxies. In this case, we exploit that the orientation of source galaxies should be random relative to the clusters, which suggest a symmetrized version of the response in the tangential frame:

$$\langle\mathsf{R}^{\rm (T)}_{\rm sel}\rangle \approx \frac{1}{2}{\rm Tr}\langle\mathsf{R}_{\rm sel}\rangle \;\;\text{where}\;\; \langle\mathsf{R}_{\rm sel}\rangle_{i,j} \approx \frac{\langle e_i\rangle^{S+} - \langle e_i\rangle^{S-}}{\Delta\gamma_j}\,. \tag{2.5}$$

In the above equation $\langle e_i\rangle^{S\pm}$ denotes the mean *un-sheared* ellipticity of galaxies when selected based on their artificially *sheared* images. Four sheared images are created by applying positive (+) and negative (–) shears of magnitude $\Delta\gamma = 0.01$ along the $\hat{e}_1$ and $\hat{e}_2$ directions separately. Errors introduced from this approximation are sub-dominant due to the already small bias associated with source galaxy selection. A detailed discussion of additional possible systematics in our specific analysis is presented in section 2.4.1.

Blinding procedure

As a precaution against unintentional confirmation bias in the scientific analyses, both weak lensing shape catalogs produced for DES Y1 had an unknown *blinding* factor in the magnitude of $\mathbf{e}$ (Zuntz et al., 2018) applied to them. This unknown factor was constrained between 0.9 and 1.1. While we made initial blinded measurements for this work, the factor was revealed as part of unblinding the cosmology results of Abbott et al. (2018).

In accordance with the practices of other DES Y1 cosmology analyses, we have further adopted a secondary layer of blinding. Specifically, we blindly transform the chains from our MCMCs to hide our in-progress results, and to prevent comparison between our cluster masses and those estimated using mass–observable relations from the literature. Chains of the parameters in the modeled lensing profiles and the mass–richness relation were unaltered after unblinding.

2.2.4 Photometric redshift catalog

In interpreting the weak gravitational lensing signal of galaxy clusters as physical mass profiles we need to employ information about the geometry of the source-lens systems by considering the relevant angular-diameter distances. To calculate these distances we rely on estimates of the overall redshift distribution of source galaxies, and also on information about the individual $P(z)$ of source galaxies.

We use the DES Y1 photometric redshifts estimated and validated by Hoyle et al. (2018) using the template-based BPZ algorithm (Benítez, 2000; Coe et al., 2006). It was found by Hoyle et al. (2018) that these photo-z estimates were modestly biased, introducing an overall multiplicative systematic correction in the recovered weak lensing profiles. We determine this correction and its systematic uncertainty in section 2.4.3.

In order to be able to correct selection effects due to the change of photo-z with shear while utilizing the highest signal-to-noise flux measurements for determining the source redshift distribution, we use two separate BPZ catalogs: one generated from METACALIBRATION-measured photometry (for selecting and weighting sources), and one from MOF (see section 2.2.1) photometry (for determining the resulting source redshift distributions). Details of this are described in the following section.

2.3 Stacked lensing measurements

2.3.1 Mass density profiles

Gravitational lensing induces distortions in the images of background galaxies, often called "sources". In the limit of weak gravitational lensing, these are characterized by the ellipticity e, which is related to the "reduced shear" via

$$\boldsymbol{g} \equiv \frac{\gamma}{1-\kappa} = \langle \mathrm{e} \rangle. \tag{2.6}$$

In the equation above, $\boldsymbol{g}$ is the reduced shear, while γ is the shear and κ is the convergence (e.g. Bartelmann & Schneider, 2001).

The gravity of a localized mass distribution, such as a galaxy cluster, induces positive shear along the tangential direction with respect to the center of the overdensity. This net tangential shear results in the stretching and preferential alignment of the images of background galaxies along the tangential direction. The magnitude of the azimuthally averaged tangential shear γ_{T}

at projected radius R can be predicted from the line-of-sight projected surface mass density Σ of the lens mass distribution by the relation

$$\gamma_{\mathrm{T}} = \frac{\overline{\Sigma}(< R) - \overline{\Sigma}(R)}{\Sigma_{\mathrm{crit}}} \equiv \frac{\Delta\Sigma(R)}{\Sigma_{\mathrm{crit}}}. \tag{2.7}$$

Here $\overline{\Sigma}(< R)$ represents the average surface mass density within projected radius R, and $\overline{\Sigma}(R)$ represents the (azimuthal) average of the surface mass density at R. For the case of reduced shear this equation holds only in linear order, therefore we account for the effect of κ in our model described in section 2.5.3.2.

The geometry of the source–lens system modulates the amplitude of the induced shear signal, and is characterized by the critical surface mass density

$$\Sigma_{\mathrm{crit}}(z_{\mathrm{s}}, z_{\mathrm{l}}) = \frac{c^2}{4\pi G} \frac{D_{\mathrm{s}}}{D_{\mathrm{l}} D_{\mathrm{ls}}} \tag{2.8}$$

in Equation 2.7. Here D_{s}, D_{l} and D_{ls} are the angular diameter distances to the source, to the lens, and between the lens and the source. Estimating the $\Delta\Sigma$ signal thus relies on robustly estimating the redshifts of the galaxy clusters and the source galaxies. The lens redshifts are the photometric redshift estimates from the redMaPPer algorithm. The statistical uncertainty on these estimates is found to be $\Delta z_{\mathrm{l}} \approx 0.01$ (Rykoff et al., 2016), which is negligible compared to other sources of error in the lensing measurement, allowing us to treat these redshifts as point estimates.

Source redshifts are also estimated from photometry, and are described by a probability distribution $p_{\mathrm{phot}}(z_{\mathrm{s}})$ for each source galaxy. We can therefore only estimate an *effective* critical surface density

$$\langle \Sigma_{\mathrm{crit}}^{-1} \rangle_{i,j} = \int \mathrm{d}z_{\mathrm{s}}\; p_{\mathrm{phot}}(z_{\mathrm{s},i})\; \Sigma_{\mathrm{crit}}^{-1}(z_{\mathrm{l},j}, z_{\mathrm{s},i})\,, \tag{2.9}$$

where i and j index the source and the lens in a lens-source pair. Note that here we choose to express the *inverse* critical surface density, which is the predicted amplitude of the lensing signal in Equation 2.7. We consistently define it as zero if $z_s \leq z_l$. For reasons of data compression, we will in fact *not* use the full integral over $p_{\mathrm{phot}}(z)$ later, but rather replace Equation 2.9 by $\Sigma_{\mathrm{crit}}^{-1}$ evaluated at a random sample of the $p_{\mathrm{phot}}(z)$.

2.3.1.1 The lensing estimator

Due to the low signal-to-noise of individual source-lens pairs we measure the stacked (mean) signal of many source galaxies around a selection of clusters.

Sheldon et al. (2004) show that the minimum variance estimator for the weak lensing signal is

$$\widetilde{\Delta\Sigma} = \frac{\sum_{j}^{\mathrm{lens}} \sum_{i}^{\mathrm{src}} w_{i,j}\; e_{\mathrm{T};\,i,j} \Big/ \langle \Sigma_{\mathrm{crit}}^{-1} \rangle_{i,j}}{\sum_{j,i} w_{i,j}}\,, \tag{2.10}$$

where the summation goes over all source–lens pairs in some radius bin and $e_{\mathrm{T};\,i,j}$ is the tangential component of the ellipticity of source i relative to lens j. The optimal weights, proportional to the inverse variance of $e_{\mathrm{T};\,i,j} / \left\langle \Sigma_{\mathrm{crit}}^{-1} \right\rangle$, are

$$w_{i,j} = \langle \Sigma_{\mathrm{crit}}^{-1} \rangle_{i,j}^{2} \Big/ \sigma_{\gamma,i}^{2} \ , \tag{2.11}$$

where $\sigma_{\gamma,i}^{2}$ is the estimate on the variance of the measured shear estimate of galaxy i relating to both the intrinsic variance of shapes and also to the uncertainty originating from shear estimation.

2.3.1.2 Practical lensing estimator

This estimator can be equivalently understood as a mean tangential ellipticity, weighted by the expected shear signal amplitude of each galaxy $\langle \Sigma_{\mathrm{crit}}^{-1} \rangle$. It is normalized by the expected signal per unit $\Delta\Sigma$, i.e. the $\langle \Sigma_{\mathrm{crit}}^{-1} \rangle$-weighted mean of the $\langle \Sigma_{\mathrm{crit}}^{-1} \rangle$. With this in mind, and including shear and selection response (see section 2.2.3), we define the estimator we use in practice as

$$\widetilde{\Delta\Sigma} \ \equiv \ \frac{\sum\limits_{j,i} \omega_{i,j} \ e_{\mathrm{T;\,i,j}}}{\sum\limits_{j,i} \omega_{i,j} \, \Sigma'^{-1}_{\mathrm{crit};i,j} \, \mathsf{R}^{\mathrm{T}}_{\gamma,i} + \left(\sum\limits_{j,i} \omega_{i,j} \, \Sigma'^{-1}_{\mathrm{crit};i,j} \right) \langle \mathsf{R}^{\mathrm{T}}_{\mathrm{sel}} \rangle} \ . \tag{2.12}$$

In the above, $\langle \mathsf{R}^{\mathrm{T}}_{\mathrm{sel}} \rangle$ is calculated via Equation 2.5 separately for source galaxies selected in each radial bin and each richness – redshift bin, where the corresponding selections were defined by the photometric redshift estimates derived from the *sheared* METACALIBRATION photometries. The small number of source galaxies at small radii introduces some noise to the estimated response, however due to the intrinsic environmental dependence of $\mathsf{R}^{\mathrm{T}}_{\mathrm{sel}}$, this cannot be readily substituted or approximated with other, less noisy quantities. By considering the expectation value

$$\langle e_{\mathrm{T;\,i,j}} \rangle = \Delta\Sigma \, \Sigma^{-1}_{\mathrm{crit};i,j} \, \mathsf{R}^{\mathrm{T}}_{i} \ , \tag{2.13}$$

it is easy to see that the definition of Equation 2.12 yields an unbiased estimate of $\Delta\Sigma$.

Equation 2.12 includes two simplifications to make calculations less computationally demanding. First, for the normalization, we replace the expectation value of $\Sigma_{\mathrm{crit}}^{-1}$ by a Monte Carlo estimate

$$\Sigma'^{-1}_{\mathrm{crit};i,j} = \Sigma_{\mathrm{crit}}(z_{\mathrm{l}_j}, z^{\mathrm{MC}}_{\mathrm{s}_i}) \, , \tag{2.14}$$

where $z^{\mathrm{MC}}_{s_i}$ is a random sample from the $p_{\mathrm{phot}}(z_s)$ distribution estimated with BPZ using MOF photometry. Second, the weights are chosen as

$$\omega_{i,j} \equiv \Sigma_{\mathrm{crit}}^{-1} \left(z_{\mathrm{l}_j}, \langle z^{\mathrm{MCAL}}_{\mathrm{s}_i} \rangle \right) \ \text{if} \ \langle z^{\mathrm{MCAL}}_{\mathrm{s}_i} \rangle > z_{\mathrm{l}_j} + \Delta z \, , \tag{2.15}$$

with $\langle z^{\mathrm{MCAL}}_{\mathrm{s}} \rangle$ being the mean redshift of the source galaxy estimated from METACALIBRATION photometry. Given the width of our photometrically estimated $p(z)$, this is close to the optimal

weight. We use a padding of $\Delta z = 0.1$ for source selection. We found that including the source weights provided by METACALIBRATION does not introduce a significant improvement in the signal-to-noise of the measurement.

The use of two different photometric estimators is necessary because when calculating the selection response, the internal photometry of the METACALIBRATION, with measurements on sheared images, must be used for all selection and weighting of sources. Hoyle et al. (2018) find this photometric redshift estimate to have a greater scatter than the default MOF photometry. We therefore opt to use the METACALIBRATION photo-z estimates only for *selecting and weighting* source-lens pairs. When normalizing the shear signal to find $\Delta\Sigma$, we utilize the MOF-based photo-z estimates.

2.3.1.3 Data vector binned in redshift and richness

In estimating the lensing signal through Equation 2.12 we utilize a modified version of the publicly available XSHEAR code[1] and the custom built XPIPE python package.[2] The core implementation of the measurement code is identical to the one used by Melchior et al. (2017).

We group the clusters into three bins in redshift: $z \in [0.2; 0.4)$, $[0.4; 0.5)$, and $[0.5; 0.65)$, as well as seven bins in richness: $\lambda \in [5; 10)$, $[10; 14)$, $[14; 20)$, $[20; 30)$, $[30; 45)$, $[45; 60)$, and $[60; \infty)$. The redshift limit $z = 0.65$ of our highest redshift corresponds roughly to the highest redshift for which the redMaPPer cluster catalog remains volume limited across the full DES Y1 survey footprint. The $\Delta\Sigma$ profiles were measured in 15 logarithmically spaced radial bins ranging from 0.03 Mpc to 30 Mpc. For our later results we will only utilize the radial range above 200 kpc. Scales below this cut are included only in our figures and for reference purposes, and are excluded from the analysis to avoid systematic effects such as obscuration, significant membership contamination, and blending. This radial binning scheme yields similar S/N across all bins. The measured shear profiles are shown in Figure 2.4.

We find a mild radial dependence in the typical value for METACALIBRATION shear response $\langle \mathsf{R}_{\gamma,\mathrm{T}} \rangle$, the asymptotic values are 0.6, 0.58 and 0.55 as a function of increasing cluster redshift. For the selection response we find an asymptotic value of $\langle \mathsf{R}_{\mathrm{sel}} \rangle \approx 0.013$, 0.014, and 0.015.

2.3.2 Covariance matrices

The $\Delta\Sigma$ profiles estimated in the previous section deviate from the true signal due to statistical uncertainties and systematic biases. We construct a description for the covariance of our data vector below and calibrate the influence of systematic effects in section 2.4.

Statistical uncertainties originate from the large intrinsic scatter in the shapes of source galaxies, the uncertainty in estimating their photometric redshifts, and due to the intrinsic variations in the properties and environments of galaxy clusters. Furthermore, our maximum radius is much larger than the 0.22 degree median separation between clusters in the catalog. This means that source galaxies are paired with multiple clusters, possibly generating covariance between between different radial ranges and/or across different cluster bins in richness and redshift.

[1] https://github.com/esheldon/xshear
[2] https://github.com/vargatn/xpipe

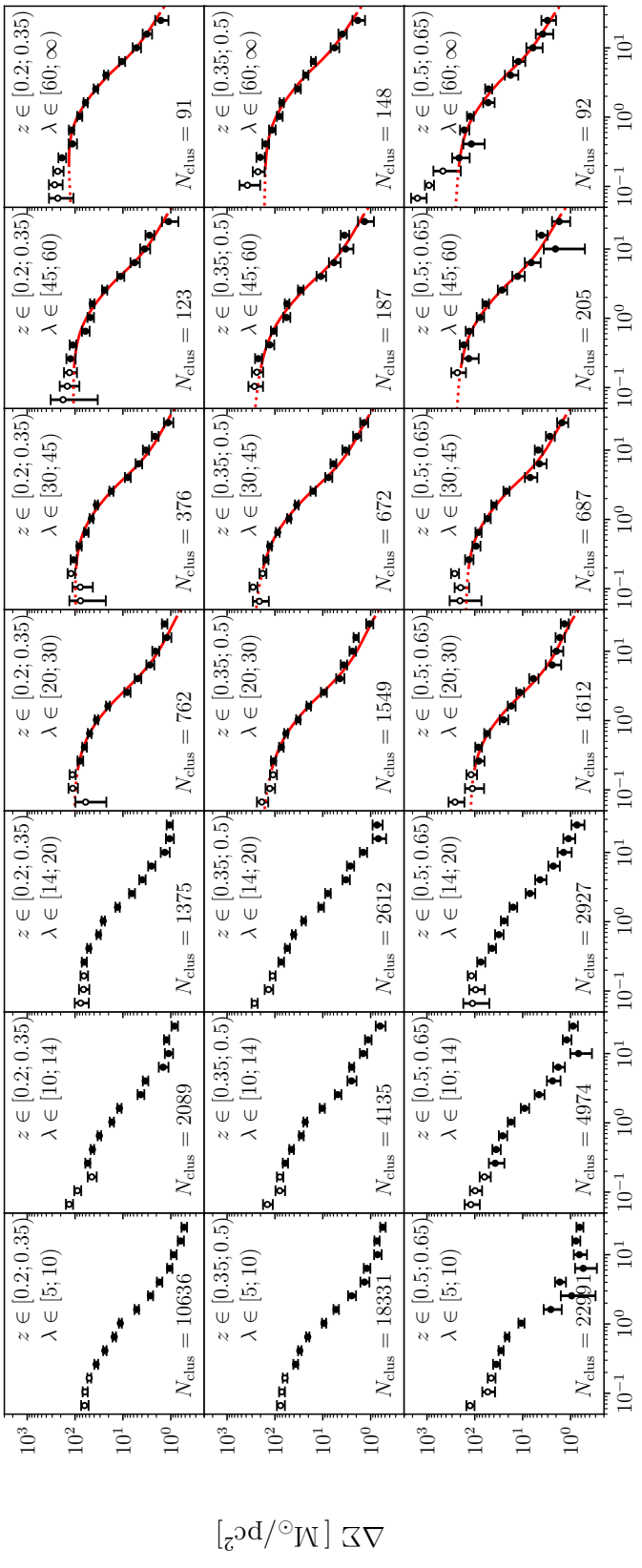

Figure 2.4: Mean $\Delta\Sigma$ for cluster subsets split in redshift $z_{\rm l}$ (increasing from top to bottom) and λ (increasing from left to right), as labeled. The error bars shown are the diagonal entries of our semi-analytic covariance matrix estimate (see section 2.3.2) for bins with $\lambda > 20$ and the jackknife estimated covariance matrix for bins with $\lambda < 20$. The best-fit model (red curve) is shown for bins with $\lambda > 20$, and includes dilution from cluster member galaxies (section 2.5.3.1) and miscentering (section 2.5.2); see section 2.5 for details. Semi-analytic covariances were not computed for stacks with $\lambda < 20$ due to the significant computational cost; the runtime scales with the number of clusters in each stack. Below 200 kpc we consider data points unreliable and therefore exclude them from our analysis; these are indicated by open symbols and dashed lines. The profiles and jackknife errors are calculated after the subtraction of the random-point shear signal (see section 2.4.1.3).

To quantify the correlation and uncertainty involved in the measurement we construct a semi-analytic model for the data covariance matrix following the framework developed by Gruen et al. (2015). Our use of a semi-analytic covariance (SAC) matrix is motivated by explicit covariance estimators exhibiting non-negligible uncertainty and possible biases, for instance from jackknife regions that are not completely independent. Both of these problems lead to a biased estimate of the precision matrix (i.e. the inverse covariance matrix), which in turn will bias the posteriors of likelihood inference (Friedrich et al., 2016).

Instead, we *predict* several key contributions of the observed covariances, namely those due to correlated and uncorrelated large scale structure, stochasticity in cluster centering, the intrinsic scatter in cluster concentrations at fixed mass, cluster ellipticity, and the scatter in the richness–mass relation of galaxy clusters. Only the shape noise contribution is estimated directly from the data, as detailed below.

While we rely on the SAC matrix estimates in the remainder of our analysis, we compare the SAC matrices to those derived using a standard jackknife method. We use jackknife (JK) resampling with $K = 100$ simply-connected spatial regions $\mathcal{R}_k$ selected via a *k-means* algorithm on the sphere.[3] The jackknife covariance is defined following Efron (1982):

$$C_{\widetilde{\Delta\Sigma}} = \frac{K-1}{K} \sum_{k}^{K} \left(\widetilde{\Delta\Sigma}_{(k)} - \widetilde{\Delta\Sigma}_{(\cdot)}\right)^{T} \cdot \left(\widetilde{\Delta\Sigma}_{(k)} - \widetilde{\Delta\Sigma}_{(\cdot)}\right) , \tag{2.16}$$

where $\widetilde{\Delta\Sigma}_{(\cdot)} = \frac{1}{K}\sum_k \widetilde{\Delta\Sigma}_{(k)}$ and $\widetilde{\Delta\Sigma}_{(k)}$ denotes the lensing signal estimated via Equation 2.12 using all lenses except those in region $\mathcal{R}_k$. Using this method, we calculate the covariance between all radial bins in a single richness and redshift bin, as well as the covariance between adjacent richness and redshift bins.

Figure 2.5 shows an example of the structure of the jackknife estimated correlation matrix between neighboring bins in richness and redshift. We find no significant correlation between richness/redshift bin and therefore treat each bin independently, even though some systematic parameters may be shared between bins.

2.3.2.1 Shape noise

The large intrinsic variations of the shapes of galaxies (*shape noise*) in the source catalog constitute a dominant source of uncertainty in lensing measurements. We estimate the covariance originating from both the random intrinsic alignments and also the stochastic positions of source galaxies. In order to do so, we make use of the measurement setup outlined in section 2.3.1.3, but each source is randomly *rotated* to create a new source catalog. We generated 1000 such independent *rotated* source catalogs, and performed the lensing measurement with each. The resulting data vectors are consistent with zero, as the random rotation washes away the imprint of the weak lensing signal. However, their scatter is indicative of the covariance due to shape noise.

[3] `https://github.com/esheldon/kmeans_radec`

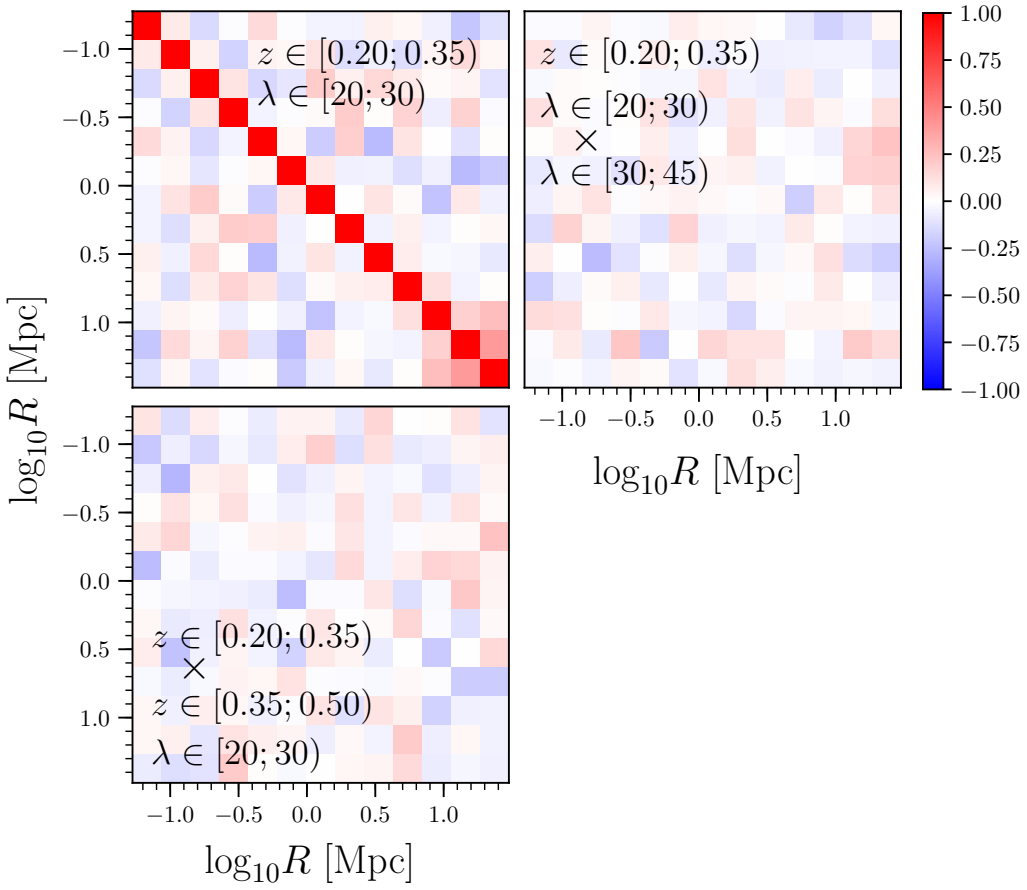

Figure 2.5: Jackknife estimated correlation matrix of $\widetilde{\Delta\Sigma}$ of a single richness-redshift selection with $\lambda \in [20; 30)$ and $z \in [0.2; 0.35)$ *(upper left panel)*. The off-diagonal blocks display the correlation matrix between the reference profile and the neighboring richness bin $\lambda \in [30; 45)$ *(upper right panel)*, and the neighboring redshift bin $z \in [0.35; 0.5)$ *(lower left panel)*

We estimate the shape noise covariance matrix for each of the 1000 realizations using the spatial jackknife scheme outlined above in section 2.3.2. The final shape noise covariance matrix estimate is obtained by averaging all 1000 of these jackknife covariance matrices. We expect this method to be less noisy compared to estimating the covariance matrix from the 1000 independent measurements of the rotated $\Delta\Sigma$ vector only.

2.3.2.2 Uncorrelated LSS

Line of sight structures which are not physically connected to the cluster leave an impact on the lensing signal. We cannot remove them from the signal, but we can estimate their expected contribution to the covariance. For an individual cluster, the covariance of the $\Delta\Sigma$ profile at radii θ_i, θ_j due to uncorrelated large scale structure (uLSS) can be written as (e.g. Schneider et al., 1998; Hoekstra, 2003; Umetsu et al., 2011)

$$C_{ij}^{\mathrm{uLSS}} = \langle \Sigma_{\mathrm{crit}}^{-1} \rangle^{-2} \int \frac{l \mathrm{d}l}{2\pi} P_\kappa(l) J_2(l\theta_i) J_2(l\theta_j) \,, \tag{2.17}$$

where P_κ is the power spectrum of the convergence, and J_2 is the Bessel function of the first kind of order 2.

Naively, one would expect that the variance of a cluster stack due to uncorrelated large scale structure to scale simply as $1/N_{\mathrm{clusters}}$. In practice, however, the positions of galaxy clusters are correlated, and the area around them overlaps on large scales. Consequently, we expect the variance due to uncorrelated structure to decrease somewhat more slowly than $1/N_{\mathrm{clusters}}$.

We estimate this source of noise by measuring random realizations of the signal due to shear fields induced by log-normal density fields with the appropriate power spectra and skewness. We calculate the latter with the perturbation theory model of Friedrich et al. (2018) for the Buzzard cosmology (DeRose et al., 2019; Wechsler et al., 2018), using the log-normal parameter κ_0 at a 10' aperture radius. As our cluster sample spans a range in redshift, a different shear field is calculated for each of the three redshift bins. This is done such that the shear fields are calculated at the lens-weighted mean source galaxy redshifts found during the initial measurement in section 2.3.1.3. We then pass these shear fields through the measurement pipeline using a spatial mask reflecting the actual source number density variations across the footprint, and estimate the covariance matrix for each realizations using 100 spatial jackknife regions for each bin in richness and redshift.

This above procedure was repeated 300 times, and the final covariance matrix due to uncorrelated LSS is taken to be the mean of the 300 jackknife covariance estimates.

2.3.2.3 Correlated LSS and halo ellipticity

Following Gruen et al. (2015), we model correlated large scale structure using a halo model approach. We assume correlated halos can be adequately described using only two parameters: the mass M of the correlated halo, and the projected distance R_{h} from the cluster. The mass distribution of the halos is assumed to follow the halo mass function, while their spatial distribution

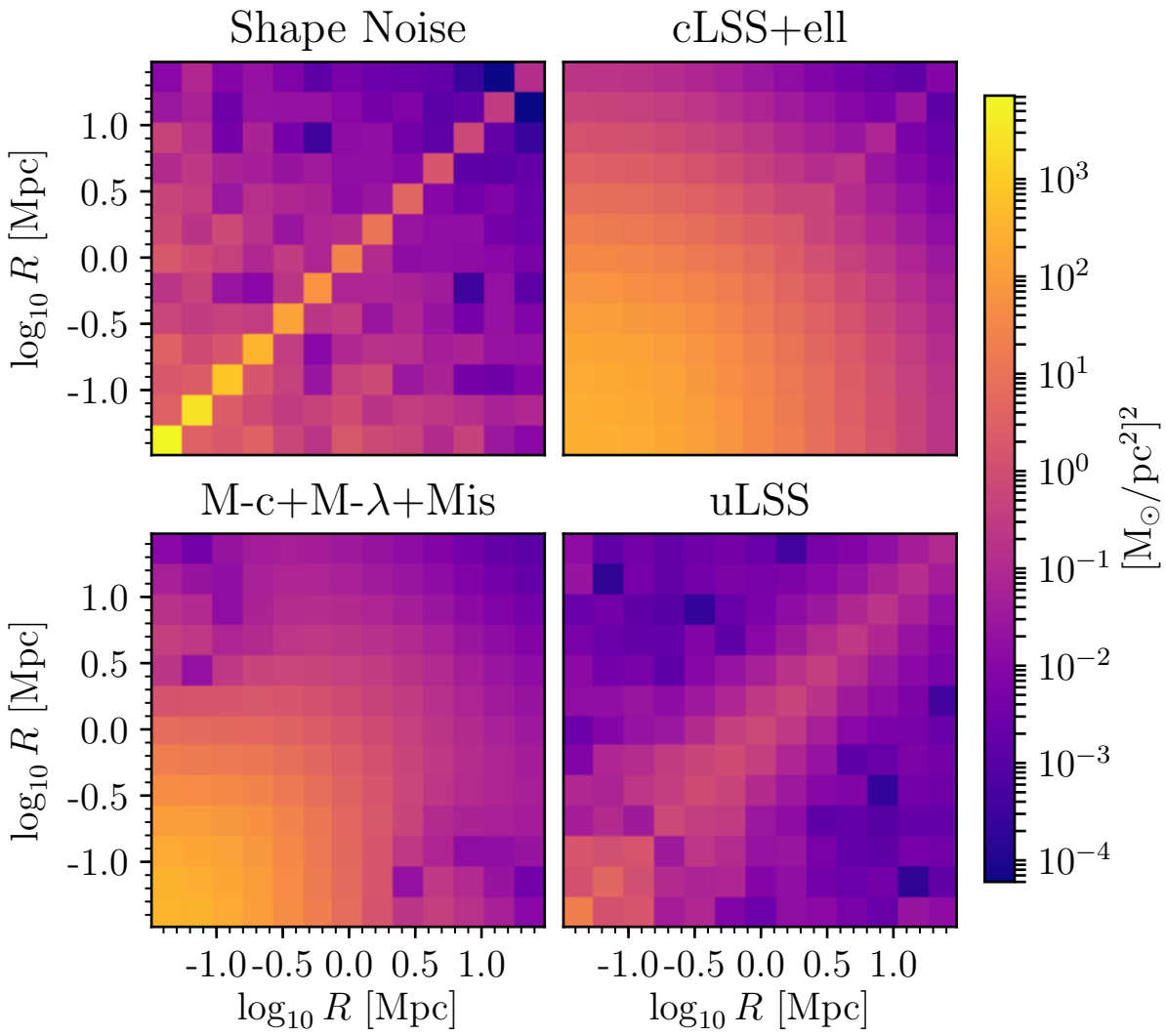

Figure 2.6: The four individual components to our Semi-Analytic Covariance (SAC) matrix. *Clockwise from top left*: Shape noise component from randomly rotating sources, correlated LSS and ellipticity component from integrating over configurations of the host cluster and its correlated halos, uncorrelated LSS from integrating over large scale structure, and finally scatter in the $M - \lambda$ relation, $M - c$ relation, and miscentering distribution. Dark colors correspond to low covariance and the colors are log scaled to show trends. Light colors are normalized to the total covariance in the SAC. See section 2.3.2 for details.

is modeled as a Poisson realization of the density field defined by the appropriate halo–cluster correlation function. That is, the excess density of halos of mass M a distance R from the halo is

$$\rho_{\rm h}(M, R_{\rm h}) = b(M_{\rm cl})b(M)w_{\rm mm}(R_{\rm h})\frac{{\rm d}n}{{\rm d}M}\,, \tag{2.18}$$

where $w_{\rm mm}$ is the projected linear correlation function at the redshift of the cluster, $b(M)$ is the halo bias, and ${\rm d}n/{\rm d}M$ is the halo mass function, or the number of halos per unit volume per unit mass (Tinker et al., 2008).

Given a model for the halo profile $\Sigma(R|M)$, the contribution of a halo at location $R_{\rm h}$ to the mean surface density Σ of the cluster in radial bin R_i is $\Sigma_i(M, R_{\rm h}) = \Sigma_{\rm misc}(R_i|M, R_{\rm h})$, where $\Sigma_{\rm misc}$ is a miscentered halo profile. Because of the Poisson-sampling assumption, the covariance matrix is generated by the mass profiles of individual halos, so that the correlated large scale structure contribution to the covariance matrix can be written as

$$C_{ij}^{\rm cLSS} = \int (2\pi R_{\rm h}{\rm d}R_{\rm h}){\rm d}M\ \rho_{\rm h}(M, R_{\rm h})\Delta\Sigma_i(M, R_{\rm h})\Delta\Sigma_j(M, R_{\rm h})\,. \tag{2.19}$$

In practice, the above predicted covariance matrix is further rescaled by a constant factor calibrated on simulations. This is meant to account for additional variance not captured by linear bias and Poisson noise, due to filamentary structure and higher order statistics in the spatial distribution of the correlated halos (e.g. the non-zero three-point function). A more detailed derivation of the above equation and its calibration is found in Gruen et al. (2015).

A very similar calculation can be made for characterizing the contribution due to halo ellipticity (to the covariance matrix – for the effect on the mean signal, see section 2.5.4.2). If $\rho_{\rm ell}(q, \mu)$ is the distribution of the halo axis ratio q and the line-of-sight orientation angle θ relative to the major axis such that $\cos\theta = \mu$, then one finds (Gruen et al., 2015)

$$C_{ij}^{\rm ell} = \int {\rm d}q{\rm d}\mu\ \rho_{\rm ell}(q, \mu)\Delta\Sigma_i\Delta\Sigma_j\,, \tag{2.20}$$

where $\Delta\Sigma_i$ is the contribution to the bin R_i under the assumption that the halo has an axis ratio q and an orientation μ.

2.3.2.4 M–c scatter, M-λ scatter and miscentering

Halos at a given mass have some intrinsic scatter in their M–λ relation. A rough estimate of the intrinsic scatter in the mass–richness (M–λ) relation is $\sim$ 25 per cent (Rozo & Rykoff, 2014; Farahi et al., 2019), and it causes an increase in the variance of stacked measurements of $\Delta\Sigma$. This scatter causes an even larger increase in the variance, since it propagates into quantities that depend directly on the mass, including the $M-c$ relation. In addition, concentration (e.g. Diemer & Kravtsov, 2015; Bhattacharya et al., 2013) and miscentering possess some intrinsic scatter from halo to halo themselves.

Scatter in the M–λ relation causes variance on all scales, since the bias $b(M)$ directly depends on the mass. By comparison, scatter in the $M - c$ relation primarily affects small scales where

the 1-halo term dominates. Similarly, some cluster centers are misidentified in our stacks, which creates additional covariance at small scales where the signal is substantially suppressed.

We modeled the combined contribution to the SAC from scatter in M–λ, scatter in concentration at fixed mass, and miscentering of individual clusters in our stacks by doing the following:

1. For each cluster in our stack, assign a mass by inverting a fiducial M–λ relation (Melchior et al., 2017) and assuming 25 per cent scatter. This is not identical to 25 per cent scatter in the M–λ relation, however this choice negligibly affects this component of the covariance matrix.

2. For each cluster, assign a concentration (including scatter) based on Diemer & Kravtsov (2015).

3. For each cluster, make a draw from our centering prior described in section 2.5.2. In other words, some fraction $f_{\rm mis}$ of clusters in the stack are miscentered, and the distribution of the amount of miscentering is given by $p(R_{\rm mis})$.

4. Calculate $\Delta\Sigma$ for each cluster and average these signals to generate a signal for the entire stack.

5. Repeat this process many times, and use these independent realizations to estimate the corresponding covariance matrix between the various radial bins.

Using Simet et al. (2017b) as our fiducial M–λ relation or using the Bhattacharya et al. (2013) mass–concentration relation had no impact on the final SAC matrix. We have also verified that using half as many realizations as our fiducial choice (1000) did not appreciably change the resulting covariance matrix. The same is true for changes in the richness scatter or miscentering model parameters within reasonable ranges.

2.3.2.5 Semi-analytic covariance matrix

Following Gruen et al. (2015), the full SAC matrix is obtained by adding each of the above contributions. The individual components described in the previous subsections are shown in Figure 2.6. Figure 2.7 demonstrates the differences between the SAC and jackknife covariance matrices. The top two panels show the correlation matrix R of the SAC and $C^{\rm JK}$ respectively, where the correlation matrix is defined via

$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}\,. \tag{2.21}$$

Similarly, to visualize the difference between C^{SA} and C^{JK} we define the residual matrix

$$Q_{ij} = \frac{C^{\rm SA}_{ij} - C^{\rm JK}_{ij}}{\sqrt{C^{\rm SA}_{ii}C^{\rm SA}_{jj}}}. \tag{2.22}$$

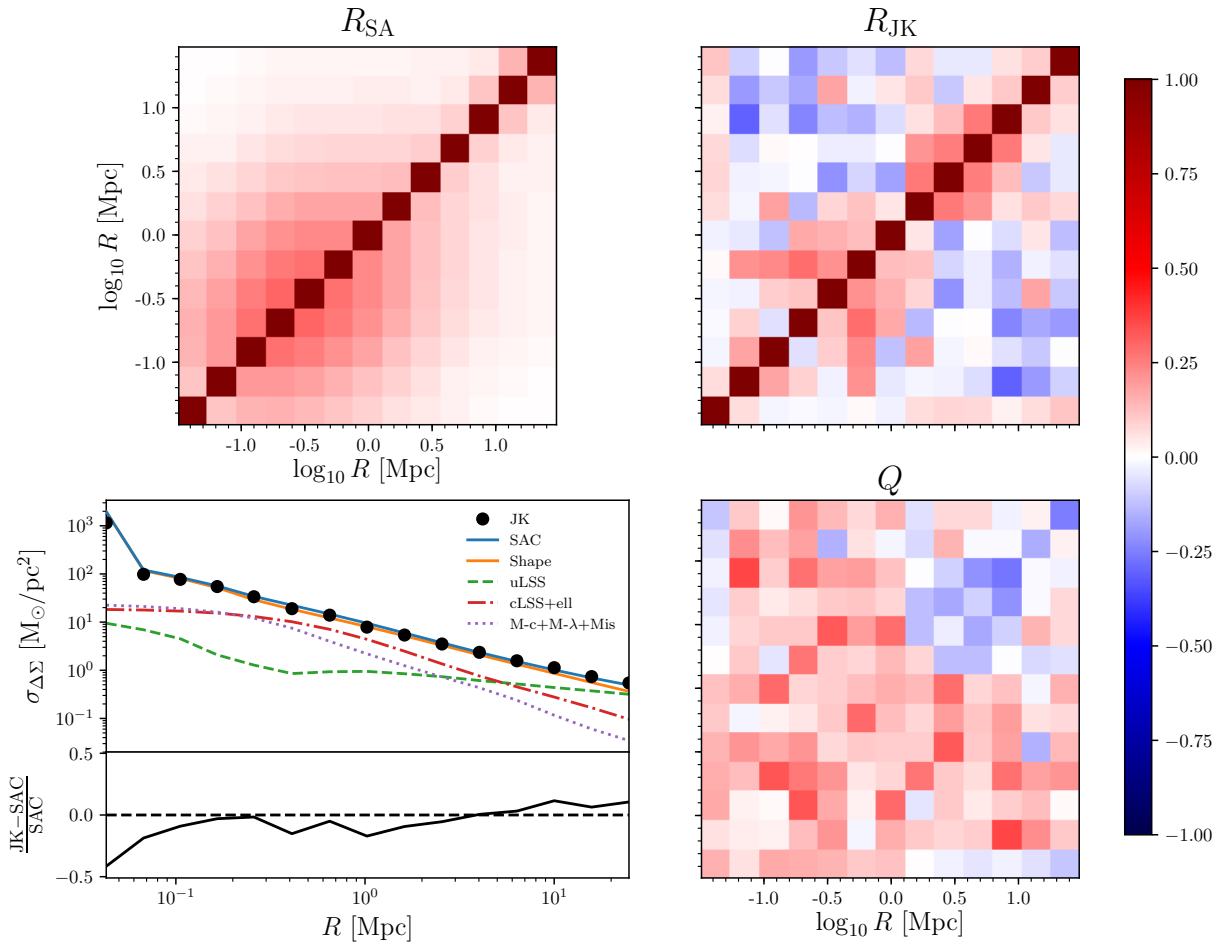

Figure 2.7: Comparison between the semi-analytic covariance matrix and the jackknife estimated covariance matrix. *Top left*: Correlation matrix of the SAC matrix. *Top-right:* Correlation matrix of the jackknife estimate. *Bottom left:* Comparison of the errors in the SAC and jackknife estimate along with the contributions to the SAC error from each individual component. The line showing the SAC errors lies almost on top of the shape noise contribution, confirming that it is the dominant source of covariance. *Bottom right:* Residual matrix Q (see Equation 2.22) that represents the difference between the SAC and jackknife covariance matrices. See section 2.3.2 for details.

We show this residual matrix in the bottom right panel of Figure 2.7. Finally, in the lower left panel we show the difference between the errors along the diagonal between the $C^{\rm JK}$ and the SAC, along with each of the contributions to the SAC; the lower panel shows the fractional difference between the diagonal entries. As expected, shape noise is the dominant contributor to the SAC matrix, with uncorrelated LSS becoming important at the largest scales. This explains why the choices we had to make in modeling the non-shape noise components did not significantly affect the resulting SAC matrix or the posteriors analysis.

Using the SACs in our analysis provides two major improvements: minimal bias from inverting the covariance matrix, and less overall noise in the off-diagonal elements which improves the mass measurement. In Melchior et al. (2017) we demonstrated that noise in the jackknife covariance matrix led to an increase of ≈30 per cent in the uncertainty of the mass of the stack. Using the SACs reduces the contribution of the covariance to the error budget by 10 per cent compared to the jackknife estimated covariance.

2.4 Systematics

2.4.1 Shear systematics

The METACALIBRATION shear catalog and the associated calibration of the source redshift distributions (Hoyle et al., 2018) passed a large number of tests performed by Zuntz et al. (2018) and Prat et al. (2018). Here we briefly enumerate the constraints on the most relevant systematics, and refer the reader to the corresponding papers for a more detailed analysis.

We parametrize the various potential biases in the dataset as:

$$g_i = (1 + m_i) g_i^{\rm tr} + \alpha^{\rm PSF} e_i^{\rm PSF} + c_i \,, \tag{2.23}$$

where $g_i^{\rm tr}$ is the true shear, while g_i is the shear estimate, and $\alpha^{\rm PSF}$ relates to the contamination from the PSF ellipticity $e_i^{\rm PSF}$.

In weak lensing surveys the three main sources of bias are commonly found to be model bias, noise bias, and selection bias (or representativeness bias). In order to account and correct for these sources of error, the METACALIBRATION algorithm performs a self-calibration on the actual data by shearing the galaxy images during the measurement, and using the thus calculated responses to correct the shear estimates. To quantify the effectiveness of this self-calibration, Zuntz et al. (2018) ran the METACALIBRATION pipeline on a set of simulated galaxy images using GALSIM (Rowe et al., 2015). The images were produced from high resolution galaxy images from the COSMOS sample, and processed to resemble the actual DES Y1 observations both in noise and PSF properties. Based on this test scenario Zuntz et al. (2018) found no significant multiplicative bias m or additive bias c present in the dataset.

Zuntz et al. (2018) further investigated the multiplicative biases due to blending of galaxy images, due to the potential leakage of stellar objects into the galaxy sample, and due to potential errors in the modeling of the PSF. They found blending as the only component with a net bias, with the other sources being consistent with zero, although contributing to the uncertainty on the

value of m. The final multiplicative bias estimates were found to be $m = [1.2 \pm 1.3] \cdot 10^{-2}$ with a 1σ Gaussian error. They found no evidence of a significant additive bias term.

Prat et al. (2018) tested for the presence of residual shear calibration biases in the DES Y1 galaxy-galaxy lensing analysis by splitting the source sample by various galaxy properties and parameters of the observational data. They showed that within the statistical uncertainty of the respective galaxy-galaxy lensing signals, and including the differences in redshift distributions induced by the splitting, no differential multiplicative biases between any of the splits are significantly detected.

In addition to the above calibrations during the construction of the shear catalog, we perform additional sanity checks relevant to stacked weak lensing measurements in the subsections below.

2.4.1.1 Second order shear bias

Due to the larger tangential shear near massive clusters, this analysis is more strongly affected by non-linear shear response than previous DES Y1 lensing analyses (see the discussion in section 9 of Sheldon & Huff 2017). This response biases cluster masses higher than they would be otherwise. To test this effect, we modify the measured $\Delta\Sigma$ profiles by adding the leading non-linear shear bias term, at third order in $\gamma_t = \Delta\Sigma \times \Sigma_{\rm crit}^{-1}$, as

$$\Delta\Sigma'_{\rm obs} = \frac{\Delta\Sigma_{\rm obs}\langle\Sigma_{\rm crit}^{-1}\rangle - \alpha_{\rm NL}\left(\Delta\Sigma_{\rm Model}\langle\Sigma_{\rm crit}^{-1}\rangle\right)^3}{\langle\Sigma_{\rm crit}^{-1}\rangle}, \tag{2.24}$$

where $\Delta\Sigma_{\rm Model}$ is the optimized model discussed in section 2.5. For the amplitude of non-linear shear bias we choose $\alpha_{\rm NL} = 0.6$ (Sheldon & Huff, 2017). We model the profile of the highest richness stack at $z \in [0.2, 0.35]$ where, for the source redshift distribution of DES Y1, this effect is strongest. The recovered mass changes by less than 1 per cent, demonstrating that our recovered mass–richness–redshift relation is robust to non-linear shear bias within our error budget.

The choice of $\alpha = 0.6$ in our test is motivated by the image simulations used in Sheldon & Huff (2017). Other simulations find a range of values of similar magnitude. Since the effect is smaller than the overall shear uncertainty, yet its calibration is uncertain, we choose not to implement a correction in our final model.

2.4.1.2 B-modes

Gravitational lensing due to localized mass distributions can only produce a net E-mode signal in the shear field, which corresponds to the tangential shear γ_t. This allows for a simple null test for the presence of systematics: any non-zero cross-shear (i.e. a non-zero B-mode) must be due to systematics. We compute the cross-shear by projecting the shears to the direction 45° from the tangential direction. We estimate the stacked B-mode signal for all richness and redshift bins, and calculate the corresponding χ^2 values using the jackknife estimate of the covariance matrix. We find $\chi^2/11 < 18/11$ for all richness bins with $\lambda > 20$, indicating that our measurement is consistent with no systematics at a $p > 0.1$ level.

2.4.1.3 Random point test

In spite of not being detected by Zuntz et al. (2018) and section 2.4.1.2, additive shear systematics may be present in the data, which could manifest as net signals visible on all radial scales. In order to test for such potential systematics we measure the lensing signal around a set of random points chosen by the redMaPPer algorithm (Rykoff et al., 2016). These points are selected via weighted random draws to mirror the distribution of DES Y1 redMaPPer clusters both in angular distribution, as well as in redshift and richness.

As additive systematics would affect the lensing profiles of galaxy clusters and random points the same way, the systematic effect can be calibrated out by subtracting the profile of random points from the profile of clusters. While we find no significant net signal around random points, we nevertheless apply this calibration, and subtract the signal of 10^5 random points from the $\widetilde{\Delta\Sigma}$ of each bin in richness and redshift. Thanks to the large number of random points used, this subtraction does not introduce significant noise to the measurement.

A motivation for subtracting the signal around random points from the measurement, regardless of the presence of systematics, is presented by Singh et al. (2017). They found that the random subtracted signal relates to the *matter over-density field* around the lenses, while the un-subtracted lensing signal traces the *matter density field*, which carries additional variance on large scales. Indeed, the precursor study of the present paper (Melchior et al., 2017) found a similar trend. We note that when constructing our SAC matrix we always apply the random point subtraction described above to ensure that our covariance matrix properly accounts for the reduced covariance that this estimator enables.

2.4.2 Correction for cluster members in the shear catalog

Due to uncertainty in photometric redshift estimates, foreground galaxies can be included in the source catalog used in our lensing measurements. So long as the ensemble redshift distribution $\mathrm{d}n/\mathrm{d}z$ of the sources is properly estimated, this is accounted for in our analysis. In the projected vicinity of galaxy clusters there is however a systematic effect biasing the naive redshift estimates of galaxies: the presence of a large cluster member population and the associated large-scale matter overdensity localized at the cluster redshift. For rich clusters, these member galaxies could make up a significant fraction of all detected galaxies in a particular line of sight. Consequently, due to intrinsic imperfections in the selection, some of these galaxies leak into the source catalog used in the weak lensing measurement. Cluster member galaxies are randomly aligned (Sifón et al., 2015), meaning their contamination results in a measured lensing signal which is biased low due to the dilution of actual source galaxies within the catalog.

It is therefore important for weak lensing studies to characterize and correct this dilution when interpreting the measurements.[4] There have been several approaches in the literature to correct for the net effect of cluster member contamination. For instance, Sheldon et al. (2004) estimated the correction factor from the transverse correlation of source galaxies around galaxy clusters, while Gruen et al. (2014) and Melchior et al. (2017) estimated the contamination rate

[4]This correction is also referred to as a *boost factor* as the measured signal should be boosted to correct for the contamination

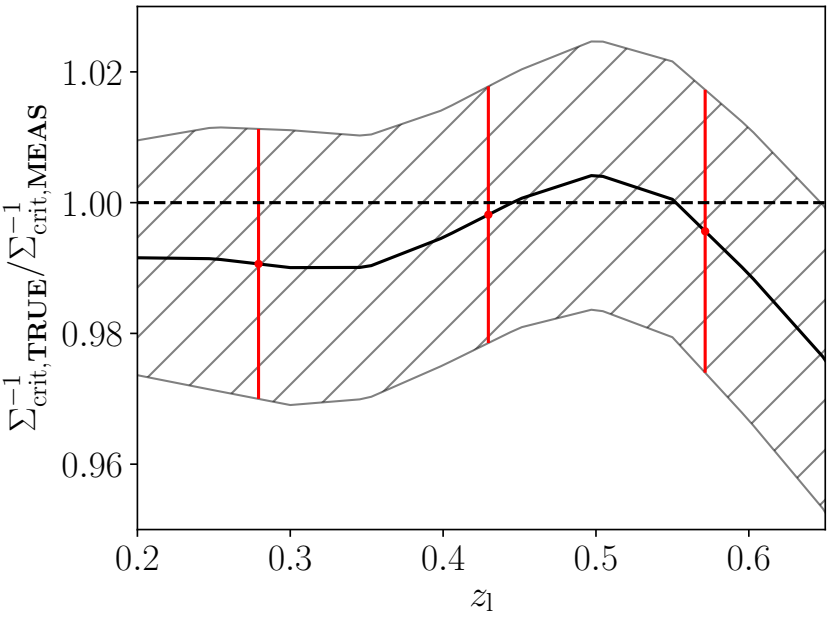

Figure 2.8: The photo-z correction factor to $\Sigma^{-1}_{\rm crit}$ as described in section 2.4.3. The gray hatched region indicates the 1σ range of the correction factor. Red points with error bars show the correction factors applied to each redshift bin.

based on the color or photometric redshift $p(z)$ information of galaxies in different radial separations from the cluster. One can also make simple color cuts (Schrabback et al., 2018b; Medezinski et al., 2010, 2018c) or photo-z cuts (Applegate et al., 2014b) on the source population to mitigate the contamination, or estimate its effect based on the increased galaxy number density around the lenses (Dietrich et al., 2019; Hoekstra et al., 2015; Simet et al., 2017a).

In this study we adopt the approach of our precursor analysis from the Science Verification data release of DES (Melchior et al., 2017), in which we make use of the estimated $p(z)$ of the source galaxy sample to calculate the cluster contamination fraction $f_{\rm cl}$ along with a corresponding covariance matrix $\mathrm{C}_{\mathrm{f_{cl}}}$ estimated via spatial jackknifing, and use this quantity to recover the contamination corrected lensing profile:

$$\widetilde{\Delta\Sigma}_{\rm corr}(R) = \frac{\widetilde{\Delta\Sigma}(R)}{1-f_{\rm cl}(R)}\,. \tag{2.25}$$

Using this $p(z)$ decomposition approach is further motivated by the complexity of the shear selection function in our analysis, which limits our ability to measure the correlation function of source galaxies. A detailed description of this method, along with validation on simulated DES-like mock observations is presented in an accompanying paper Varga et al. (2019).

2.4.3 Photometric redshift systematics

The redshift distribution of our selected source galaxies was estimated using BPZ (Benítez, 2000) in the implementation of Hoyle et al. (2018). In BPZ or similar photometric redshift estimation procedures, one assumes a variety of galaxy spectral energy distribution (SED) templates and priors for the relative abundance of galaxies as a function of luminosity and redshift. Any deviation from these assumptions in the DES source galaxy sample can cause biases in photometric redshift estimates which must be calibrated.

For the cosmology analyses of the lensing two-point functions (Troxel et al., 2017; Dark Energy Survey Collaboration, 2016), this calibration was performed in two independent ways, and with consistent results: by the redshift distributions of samples of galaxies with high-quality 30-band photo-zs from COSMOS, matched to DES lensing source galaxies (Hoyle et al., 2018), and by the clustering of lensing source galaxies with redMaGiC (Rozo et al., 2016) galaxies as a function of the redshift of the latter (Davis et al., 2018; Gatti et al., 2018).

For this work, we adapt the COSMOS calibration of Hoyle et al. (2018) to estimate the bias of our $\Delta\Sigma$ measurements, and the uncertainty in that bias. To this end, we select and weight galaxies from COSMOS in the same manner as for our measurements of the cluster $\Delta\Sigma$ profiles.

Following Hoyle et al. (2018, their section 4.1), we randomly sample 200,000 galaxies in our data and match them to COSMOS galaxies according to their flux in each band and their intrinsic size. From this *COSMOS resampling*, we select and weight galaxies as per section 2.3.1.2 and section 2.3.1.3. From the COSMOS 30-band we calculate the weighted mean true $\Sigma'^{-1}_{\rm crit,TRUE}$. From noisy MOF *griz* BPZ redshift distribution samples we get a mean $\Sigma'^{-1}_{\rm crit,MEAS}$ that relates the weighted mean tangential shear to the $\Delta\Sigma$ profile. As in the denominator of Equation 2.12, we use a weight $\omega \times \mathsf{R}$ for the means. Because the source selection, ω weight, and $\Sigma'_{\rm crit}$ depends on lens redshift, we repeat this exercise for the range of cluster redshifts sampled by our catalog, $z_l = 0.2 \ldots 0.65$. A bias in $\Sigma'^{-1}_{\rm crit}$ translates directly into a multiplicative bias in $\Delta\Sigma$.

We estimate four sources of uncertainty in the calibration of photometric redshift distributions (see Hoyle et al., 2018, their sections 4.2-4.5): 1) an uncertainty due to cosmic variance from the relative scatter of average $\Sigma'^{-1}_{\rm crit,TRUE}$ in the resampling of the 368 simulated COSMOS footprints, to which we add the (subdominant) statistical uncertainty due to the limited sample size from bootstrap resamplings in quadrature; 2) an uncertainty due to photometric zeropoint offsets from realizations of photometric zeropoint calibration offsets; 3) an uncertainty due to the morphology matching, which we estimate as half the difference between the estimated $\Sigma'^{-1}_{\rm crit,TRUE}$ of the sample with size+flux matching and that obtained without the size matching; and 4) a systematic uncertainty of the matching algorithm by a comparison between the fiducial $\Sigma'^{-1}_{\rm crit,TRUE}$ value and that of the aforementioned 368 resampled simulated COSMOS fields. Effects 1, 3, and 4 contribute to the systematic uncertainties with similar size, while effect 2 is smaller but not quite negligible.

We define our model for the bias as

$$\frac{\Sigma'^{-1}_{\rm crit,MEAS}}{\Sigma'^{-1}_{\rm crit,TRUE}} \equiv 1 + \delta\,, \tag{2.26}$$

with the mean value given from the COSMOS analysis, and an uncertainty due to the four effects mentioned above. This ratio depends on lens redshift through the selection/weighting of sources and the source redshift dependence of photo-z bias. It is plotted across the entire lens redshift range in Figure 2.8. The red points show the ratio at the mean redshifts of the bins used in our analysis. This multiplicative factor is fully degenerate with shear systematics (see section 2.4.1) and is assumed to be correlated across redshift bins. δ is incorporated in our analysis by a prior that varies between each stack. The variation between richness bins is small compared to the variation across cluster redshift bins.

$$\delta = \begin{cases} 0.009 \pm 0.021 & \text{for } z \in [0.2, 0.35] \\ 0.002 \pm 0.020 & \text{for } z \in [0.35, 0.5] \\ 0.004 \pm 0.022 & \text{for } z \in [0.5, 0.65]. \end{cases} \tag{2.27}$$

This prior is combined with the prior on m and included in the final likelihood as described in section 2.5.3.3.

2.5 The stacked lensing signal

2.5.1 Surface density model

Our surface density model remains unchanged from Melchior et al. (2017). The lensing signal is given by Equation 2.7. The quantities $\overline{\Sigma}(R)$ and $\overline{\Sigma}(< R)$ are given by

$$\overline{\Sigma}(R) = \int_{-\infty}^{+\infty} \mathrm{d}\chi \, \Delta\rho \left(\sqrt{R^2 + \chi^2} \right) , \tag{2.28}$$

where R is the projected separation and χ the separation along the line of sight in comoving units and

$$\overline{\Sigma}(< R) = \frac{2}{R^2} \int_0^R \mathrm{d}R' \, R' \overline{\Sigma}(R') \, . \tag{2.29}$$

If the shear signal is caused by halos of mass M, the average excess three dimensional matter density is given by

$$\Delta\rho(r) = \rho(r) - \rho_\mathrm{m} = \rho_\mathrm{m} \xi_\mathrm{hm}(r \,|\, M) \, , \tag{2.30}$$

where $\rho_\mathrm{m} = \Omega_\mathrm{m} \rho_\mathrm{crit} (1 + z)^3$ is the mean matter density in physical units at the redshift of the sample, ρ_crit is the critical density at redshift zero, and $\xi_\mathrm{hm}(r \,|\, M)$ is the halo–matter correlation function at the halo redshift.

At small scales ξ_hm is dominated by the so-called "1-halo" term while at large scales it is dominated by the "2-halo" term. We use the Zu et al. (2014) update to the Hayashi & White (2008) model of ξ_hm. This is

$$\xi_\mathrm{hm}(r \,|\, M) = \max \left\{ \xi_\mathrm{1h}(r \,|\, M), \xi_\mathrm{2h}(r \,|\, M) \right\} \, . \tag{2.31}$$

For the 1-halo term we use a Navarro et al. (1996, hereafter NFW) density profile $\rho_{\rm NFW}(r\,|\,M)$

$$\xi_{\rm 1h}(r\,|\,M,c) = \frac{\rho_{\rm NFW}(r\,|\,M,c)}{\rho_m} - 1\,, \tag{2.32}$$

where

$$\rho_{\rm NFW}(r\,|\,M,c) = \frac{\Omega_{\rm m}\rho_{\rm crit}\delta_c}{(r/r_s)\,(1+r/r_s)^2}\,. \tag{2.33}$$

The concentration $c = r_{200m}/r_s$ is left as a free parameter, with a flat prior as per Table 2.1. This differs from the analysis in Melchior et al. (2017), in which we forced the halo concentration to follow the concentration–mass relation of Diemer & Kravtsov (2015).

For the two-halo term $\xi_{\rm 2h}(r\,|\,M)$ we use the non-linear matter correlation function $\xi_{\rm nl}$ scaled by the halo bias $b(M)$ of Tinker et al. (2008) as

$$\xi_{\rm 2h}(r\,|\,M) = b(M)\xi_{\rm nl}(r)\,. \tag{2.34}$$

$\xi_{\rm nl}$ is the 3D Fourier transform of the non-linear power spectrum $P_{\rm nl}$ (Smith et al., 2003; Takahashi et al., 2012), given by

$$\xi_{\rm nl}(r) = \int_0^\infty \frac{{\rm d}k}{k}\,\frac{k^3 P_{\rm nl}(k)}{2\pi^2} j_0(kr)\,, \tag{2.35}$$

where $j_0(z)$ is the 0th spherical Bessel function of the first kind. The power spectrum is computed using the CLASS code[5] (Lesgourgues, 2011; Blas et al., 2011). We repeated our analysis using the linear matter correlation function $\xi_{\rm lin}$ and found nearly identical results as discussed later in section 2.6.3.1.

2.5.2 Miscentering correction

We have thus far assumed that we can measure the stacked shear profile of clusters relative to the "center" of the halo as defined in an N-body simulation. Our simulations use the spherical overdensity algorithm ROCKSTAR as implemented in Behroozi et al. (2013). If cluster centers are not properly identified, or are "miscentered", then the observed weak lensing signal in annuli around these clusters will be suppressed. As in Melchior et al. (2017), we model the recovered weak lensing signal as a weighted sum of two independent contributions: a contribution $\Delta\Sigma_{\rm cen}$ from properly centered clusters, and a contribution $\Delta\Sigma_{\rm mis}$ from miscentered clusters,

$$\Delta\Sigma_{\rm model} = (1-f_{\rm mis})\Delta\Sigma_{\rm cen} + f_{\rm mis}\Delta\Sigma_{\rm mis}\,. \tag{2.36}$$

$\Delta\Sigma_{\rm cen}$ is given by Equation 2.7. When a cluster is miscentered by some radial offset $R_{\rm mis}$, the corresponding azimuthally averaged surface mass density is (e.g. Yang et al. (2006); Johnston et al. (2007))

$$\overline{\Sigma}_{\rm mis}(R\,|\,R_{\rm mis}) = \int_0^{2\pi} \frac{{\rm d}\theta}{2\pi}\,\overline{\Sigma}\left(\sqrt{R^2 + R_{\rm mis}^2 + 2RR_{\rm mis}\cos\theta}\right)\,. \tag{2.37}$$

[5] http://class-code.net/

Letting $p(R_{\rm mis})$ be the distribution of radial offsets for miscentered clusters, the corresponding mean miscentered profile $\overline{\Sigma}_{\rm mis}$ is

$$\overline{\Sigma}_{\rm mis}(R) = \int {\rm d}R_{\rm mis}\, p(R_{\rm mis})\overline{\Sigma}_{\rm mis}(R\,|\,R_{\rm mis})\,. \tag{2.38}$$

It is this quantity that we use to model the miscentered profile term in Equation 2.36.

Zhang et al. (2019b); von der Linden et al. (2018) measure the centering fraction and centering distribution of redMaPPer clusters by comparing the reported centers to those derived from high resolution X-ray data (where available). Here, we summarize their findings. The miscentering distribution $p(R_{\rm mis})$ has the form

$$p(R_{\rm mis}) = \frac{R_{\rm mis}}{(\tau R_\lambda)^2}\exp\left(-\frac{R_{\rm mis}}{\tau R_\lambda}\right) \tag{2.39}$$

where R_λ is the cluster radius assigned by redMaPPer, and $\tau = R_{\rm mis}/R_\lambda$. Note that this is a Gamma distribution instead of the Rayleigh distribution used in, e.g. Simet et al. (2017b); Melchior et al. (2017). This model choice is justified in Zhang et al. (2019b); von der Linden et al. (2018). We use a combination of the posteriors of Zhang et al. (2019b); von der Linden et al. (2018) to set the priors of the miscentering parameters $f_{\rm mis}$ and τ, as detailed in Table 2.1. The prior uncertainties conservatively encompass the spread in best fitting values and the confidence intervals of both works. It corresponds to a miscentering fraction $f_{\rm mis} = 0.25 \pm 0.08$, that is, roughly ≈ 75 per cent of the redMaPPer clusters are correctly centered. Because the variation in R_λ within each richness bin is mild, we ignore variations in R_λ across the bin, and set R_λ to the radius of a cluster whose richness is equal to the mean richness of the clusters in the bin. We have explicitly verified that if use the median rather than the mean cluster richness, the difference between our predicted profiles is insignificant relative to our statistical errors.

2.5.3 Multiplicative corrections

Multiplicative corrections to our model include boost factors, reduced shear, and shear+photo-z biases. These adjust our model according to

$$\Delta\Sigma_{\rm full}(R) = \frac{\mathcal{A}_m\mathcal{G}(R)}{\mathcal{B}(R)}\Delta\Sigma_{\rm model}\,. \tag{2.40}$$

In this equation, $\mathcal{A}_m$ is the multiplicative correction due to shear and photometric redshift biases, $\mathcal{G}(R)$ is the multiplicative correction due to using reduced shear, and $\mathcal{B}(R)$ is the boost factor correction.

2.5.3.1 Boost factor model

In section section 2.4.2, we discussed how membership dilution biases the recovered weak lensing profile by a factor $1 - f_{\rm cl}$. This factor is known as a boost factor because to correct for it in the lensing profile, one would increase the measured signal. We decided to not apply this

factor to our data, and instead dilute the theoretical profile. To marginalize over the statistical uncertainty in our boost factor measurements, we parameterize the boost factor $\mathcal{B} \equiv (1 - f_{\rm cl})^{-1}$ by constructing a model for the cluster-member contamination using a two component (B_0 and R_s) NFW profile:

$$\mathcal{B}_{\rm model}(R) = 1 + B_0 \frac{1 - F(x)}{x^2 - 1}\,, \tag{2.41}$$

where $x = R/R_s$, and

$$F(x) = \begin{cases} \frac{\tan^{-1}\sqrt{x^2-1}}{\sqrt{x^2-1}} & : x > 1 \\ 1 & : x = 1 \\ \frac{\tanh^{-1}\sqrt{1-x^2}}{\sqrt{1-x^2}} & : x < 1 \end{cases}\,. \tag{2.42}$$

We fit the boost factors measured in each bin along with the lensing profile of that bin. This introduces two additional parameters in our model, B_0 and R_s, for each richness and redshift bin.

2.5.3.2 Reduced Shear

We account for the fact that we measure the reduced shear g rather than true shear γ, as seen in Equation 2.6, by multiplying our lensing model by

$$\mathcal{G}(R) = \frac{1}{1 - \kappa} = \frac{1}{1 - \Sigma(R)\Sigma_{\rm crit}^{-1}}\,. \tag{2.43}$$

Here, $\Sigma_{\rm crit}^{-1}$ is the same as that in 2.4.3 and $\Sigma(R)$ is

$$\Sigma(R) = (1 - f_{\rm mis})\Sigma_{\rm cen} + f_{\rm mis}\Sigma_{\rm mis}\,, \tag{2.44}$$

where $\Sigma_{\rm cen}$ is given by Equation 2.28 and $\Sigma_{\rm mis}$ is given by Equation 2.38. This adjustment has a negligible effect on our results, and introduces no new free parameters to our analysis.

2.5.3.3 Shear+photo-z bias

The factor $\mathcal{A}_m = 1 + m + \delta$ combines the effects of shear (m, section 2.4.1) and photo-z (δ, section 2.4.3) systematic uncertainties. Zuntz et al. (2018) found a shear calibration of $m = 0.012 \pm 0.013$. The photo-z bias comes from Hoyle et al. (2018) and varies between cluster stacks.

Since both m and δ are assigned Gaussian priors, the width of the prior on $\mathcal{A}_m$ is obtained by adding the widths of the priors on m and δ in quadrature. We arrive at

$$\mathcal{A}_m = \begin{cases} 1.021 \pm 0.025 & \text{for } z \in [0.2, 0.35] \\ 1.014 \pm 0.024 & \text{for } z \in [0.35, 0.5] \\ 1.016 \pm 0.025 & \text{for } z \in [0.5, 0.65]. \end{cases} \tag{2.45}$$

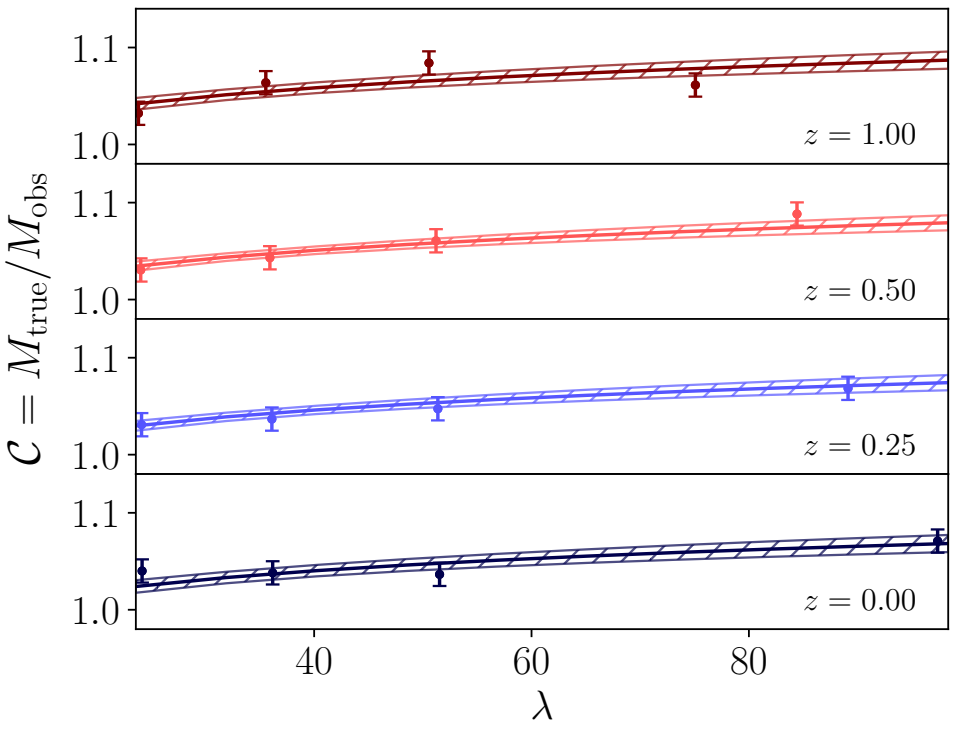

Figure 2.9: The mass calibration $\mathcal{C} = M_{\rm true}/M_{\rm obs}$ from adopting our model of the correlation function in Equation 2.31 as a function of λ and redshift. The solid line and hatched region are the best fit model and 1σ uncertainty of the calibration. Error bars on the measured calibrations are the fitted intrinsic scatter $\sigma_{\mathcal{C}}$.

The typical systematic uncertainty of ≈ 2.5 per cent represents a significant improvement over the typical systematic uncertainty of ≈ 3.8 per cent we achieved in Melchior et al. (2017). This dramatic improvement in accuracy is primarily driven by the improved shear calibration achieved in the DES Y1 data with METACALIBRATION.

For the following data releases of DES, we anticipate that improvements in the treatment of blended objects can further reduce the multiplicative shear bias. This implies that uncertainties in the calibration of photometric redshift estimates will likely be our dominant measurement related systematic. Significant improvements on this will require either extended calibration data sets or a hierarchical treatment that uses survey data to inform redshift estimation consistently.

2.5.4 Stacked mass corrections

We expect the masses we measure in section 2.5.5 to be biased with respect to the true mean mass of the stacks. This bias arises from two sources: our model presented above is not a true description of cluster lensing profiles, and effects due to triaxiality and projection. We account for both sources of bias by calculating a correction $\mathcal{C}$ applied to the expected mass of the stack $\mathcal{M}_{\rm true} = \mathcal{C}\langle M \rangle$, as detailed in the section below. This is applied after the lens modeling is com-

plete, but before modeling the mass–observable relation from our stacked masses in section 2.6.

2.5.4.1 Modeling systematics

The model presented above for $\Delta\Sigma$ is not perfect; our analytic model for the halo-mass correlation function in Equation 2.31 does not match density profiles in simulations (Melchior et al., 2017; Murata et al., 2017), in particular in the transition between the 1-halo and 2-halo regimes. In lieu of a fully calibrated model, we correct for any bias imparted by our choice of model by using our likelihood analysis to estimate halo masses of synthetic data generated from N-body simulations. The halos are drawn from an N-body simulation of a flat ΛCDM cosmology run with GADGET (Springel, 2005). The simulation uses 1400^3 particles in a box with 1050 h^{-1}Mpc on a side with periodic boundary conditions and for softening of 20 h^{-1}kpc. The simulation was run with the cosmology $\Omega_m = 0.318$, $h = 0.6704$, $\Omega_b = 0.049$, $\tau = 0.08$, $n_s = 0.962$, and $\sigma_8 = 0.835$. Halos of mass 10^{13} $h^{-1}\mathrm{M}_\odot$ are resolved with 100 particles. We discard all information below 5 softening lengths, and verified that the choice of extrapolation scheme for describing the correlation function below this scale does not impact our results. Halos were defined using a spherical overdensity mass definition of 200 times the background density and were identified with the ROCKSTAR halo finder (Behroozi et al., 2013).

The simulation is used to construct the synthetic $\Delta\Sigma$ profiles of galaxy clusters at four different snapshots: $z \in [0, 0.25, 0.5, 1]$. There were ∼420,000 halos at $z = 1$ and ∼830,000 halos at $z = 0$. We used snapshots instead of lightcones for two main reasons: we wanted to maximize the number of halos we had available to perform the calibration, and we found that the synthetic profiles to only weakly depend on redshift. Instead of splitting halos into mass subsets as in Melchior et al. (2017), we assigned a richness to each halo by inverting the mass–richness relation of Melchior et al. (2017) and adding 25 per cent scatter. We then grouped our halos into richness subsets identical to how we grouped our clusters. For each of these halo subsets we measured the halo-matter correlation function with the Landy & Szalay (1993) estimator as implemented in CORRFUNC[6] (Sinha & Garrison, 2017). We numerically integrate the halo-matter correlation function to obtain the $\Delta\Sigma$ profile as described in section 2.5.1.

This $\Delta\Sigma$ profile contains none of the systematics that exist in the real data. To incorporate them, we modified this profile with the multiplicative corrections described in section 2.5.3 and miscentering corrections in section 2.5.2. We took the central values of our priors in Table 2.1 as well as values for B_0 and R_s from modeling the boost factors independently and modified the simulated $\Delta\Sigma$ profile according to Equation 2.47. The observed mass M_{obs} for this simulated profile was obtained by using the same pipeline that we apply on the real data. When evaluating the likelihood in Equation 2.50, we used the semi-analytic covariance matrix corresponding to the nearest cluster subset in redshift.

Denoting M_{true} as the mean mass of the halos in the simulated stack, the calibration for each simulated profile is seen in Figure 2.9. The calibration $\mathcal{C} = M_{\mathrm{true}}/M_{\mathrm{obs}}$ was then modeled as a

[6] https://github.com/manodeep/Corrfunc

function of the mean richness of the simulated stack $\bar{\lambda}$ in the form

$$\mathcal{C}(\bar{\lambda}, z) = C_0 \left(\frac{\bar{\lambda}}{\lambda_0}\right)^{\alpha} \left(\frac{1+z}{1+z_0}\right)^{\beta} . \tag{2.46}$$

The free parameters in our fit are C_0, α, and β with pivot values at $z_0 = 0.5$ and λ_0=30, as well as the intrinsic scatter $\sigma_{\mathcal{C}}$ of the calibration. The mean model bias for our cluster stacks is ≈ 4 per cent with $C_0 = 1.042 \pm 0.004$, $\alpha = 0.03 \pm 0.006$, and $\beta = 0.025 \pm 0.012$ as well as $\sigma_{\mathcal{C}} = 0.01$. We repeated this process while assuming different amounts of intrinsic scatter in the M–λ relation from 10 per cent up to 45 per cent, as well as with no intrinsic scatter which is equivalent to the treatment of Melchior et al. (2017). We found that the amount of model bias increased with scatter in the M–λ relation. The model bias from Melchior et al. (2017) was recovered when no intrinsic scatter was present and using covariance matrices from that analysis.

We incorporated the dependence of the calibration on the intrinsic scatter in the M–λ relation as follows. We took the calibration described above at 25 per cent scatter to be our fiducial model as estimated in Rozo & Rykoff (2014). In addition to the covariance between C_0, α, and β, we add additional uncorrelated uncertainty to each of these terms equal to half of the difference between the mean values obtained for these parameters assuming 15 per cent and 45 per cent scatter. This increased the variance of all three parameters C_0, α and β slightly. As discussed further in 2.8, the calibration contributed 0.73 per cent to the overall systematic uncertainty on the normalization of the M–λ relation.

2.5.4.2 Triaxiality and projection effects

Photometric cluster selection preferentially selects halos that are oriented with their major axis along the line of sight. Similarly, cluster selection is affected by other objects along the line of sight, which increases both the observed cluster richness and the recovered lensing mass. These two effects have been studied closely elsewhere (White et al., 2011; Angulo et al., 2012; Noh & Cohn, 2012; Dietrich et al., 2014), and have competing effects on the recovered cluster masses. Dietrich et al. (2014) determined that triaxiality leads to an overestimation of the weak lensing mass and requires a correction factor of 0.96 ± 0.02, while Simet et al. (2017b) argued projection effects require that the recovered masses be multiplied by a factor of 1.02 ± 0.02. Together, triaxiality and projection effects modify the recovered weak lensing masses by a multiplicative factor of 0.98 ± 0.03. Our treatment is identical to that of Melchior et al. (2017), where additional details are provided. Although these two effects mildly depend on richness and redshift, we assume them to be constant in this analysis. We show the cumulative effect in Table 2.6. For reference, we have estimated the number of galaxy clusters that have another *cluster* within a 500 kpc radius along the line of sight. The number of such cases with $\lambda \geq 20$ is about 30, or 0.4 per cent of our sample, and thus negligible.

These effects as well as the correction for model bias are applied to the masses after fitting the lensing and boost factor data as described in section 2.5.5, but before modeling the M–λ relation in section 2.6.

2.5.5 The complete likelihood

The full model of the weak lensing profile is

$$\Delta\Sigma = \frac{\mathcal{A}_m \mathcal{G}(R)}{\mathcal{B}(r)} \left[(1 - f_{\rm mis})\Delta\Sigma_{\rm cen} + f_{\rm mis}\Delta\Sigma_{\rm mis}\right] . \tag{2.47}$$

Written this way the model includes the multiplicative bias $\mathcal{A}_m$, the boost factor $\mathcal{B}(r)$, the reduced shear correction $\mathcal{G}(R)$, and miscentering effects f_{mis} and $\Delta\Sigma_{\rm mis}$. Using the semi-analytic covariance matrices $\mathrm{C}_{\Delta\Sigma}$ described in section 2.3.2.5, the log-likelihood of the kth $\Delta\Sigma$ profile is

$$\ln \mathcal{L}(\Delta\Sigma_k \,|\, M_k, c, \mathcal{A}_m, R_{\rm mis}, f_{\rm mis}, B_0, R_s) \propto -\frac{1}{2}\mathbf{D}_k^T \mathrm{C}_{\Delta\Sigma}^{-1} \mathbf{D}_k \tag{2.48}$$

where $\mathbf{D} = (\widetilde{\Delta\Sigma} - \Delta\Sigma)$ and $\widetilde{\Delta\Sigma}$ is the measurement from Equation 2.12.

The boost factor covariance matrix $\mathrm{C}_{\rm f_{cl}}$ is estimated from jackknifing. With this the corresponding log-likelihood of the measured $f_{{\rm cl},k}$ in cluster subset k given the parameters in Equation 3.21 is

$$\ln \mathcal{L}(f_{{\rm cl},k} \,|\, B_0, R_s) \propto -\frac{1}{2}\mathbf{B}_k^T \mathrm{C}_{f_{\rm cl}}^{-1} \mathbf{B}_k \tag{2.49}$$

where $\mathbf{B}_k = (\mathcal{B} - \mathcal{B}_{\rm model})_k$. Each boost factor $\mathcal{B}_k$ is fit in conjunction with the associated lensing profile for the kth subset, allowing us to account for any degeneracies between the parameters in the respective models.

The total log-likelihood for a single cluster subset is

$$\begin{aligned} \ln \mathcal{L}_k = &\ln \mathcal{L}(\Delta\Sigma_k \,|\, M_k, c, \mathcal{A}_m, R_{\rm mis}, f_{\rm mis}, B_0, R_s) + \\ &\ln \mathcal{L}(f_{{\rm cl},k} \,|\, B_0, R_s) . \end{aligned} \tag{2.50}$$

Our goal is to constrain masses of independent cluster subsets M_k and boost factor parameters. Constraints on the latter are informed by both their effect on the $\Delta\Sigma$ profile as well as independent measurements of $f_{\rm cl}$. The weak lensing and boost factor profiles are fit simultaneously, but each cluster subset is fit independently of other subsets.

2.5.6 Stacked cluster masses

A complete list of the model parameters describing each cluster stack as well as their corresponding priors are summarized in Table 2.1. The likelihood is sampled using the package *emcee*[7] (Foreman-Mackey et al., 2013), which enables a parallelized exploration of the parameter space. We use 32 walkers with 10000 steps each, discarding the first 1000 steps of each walker as burn-in. We checked the convergence with independent runs of 5000 steps per walker that yielded identical results. After 14 steps the chains of single walkers become uncorrelated (with a correlation coefficient $|r| < 0.1$). This is much shorter than the total length of the chain. As a result the number of independent draws between all walkers is ≈ 20500.

[7] `http://dan.iel.fm/emcee`

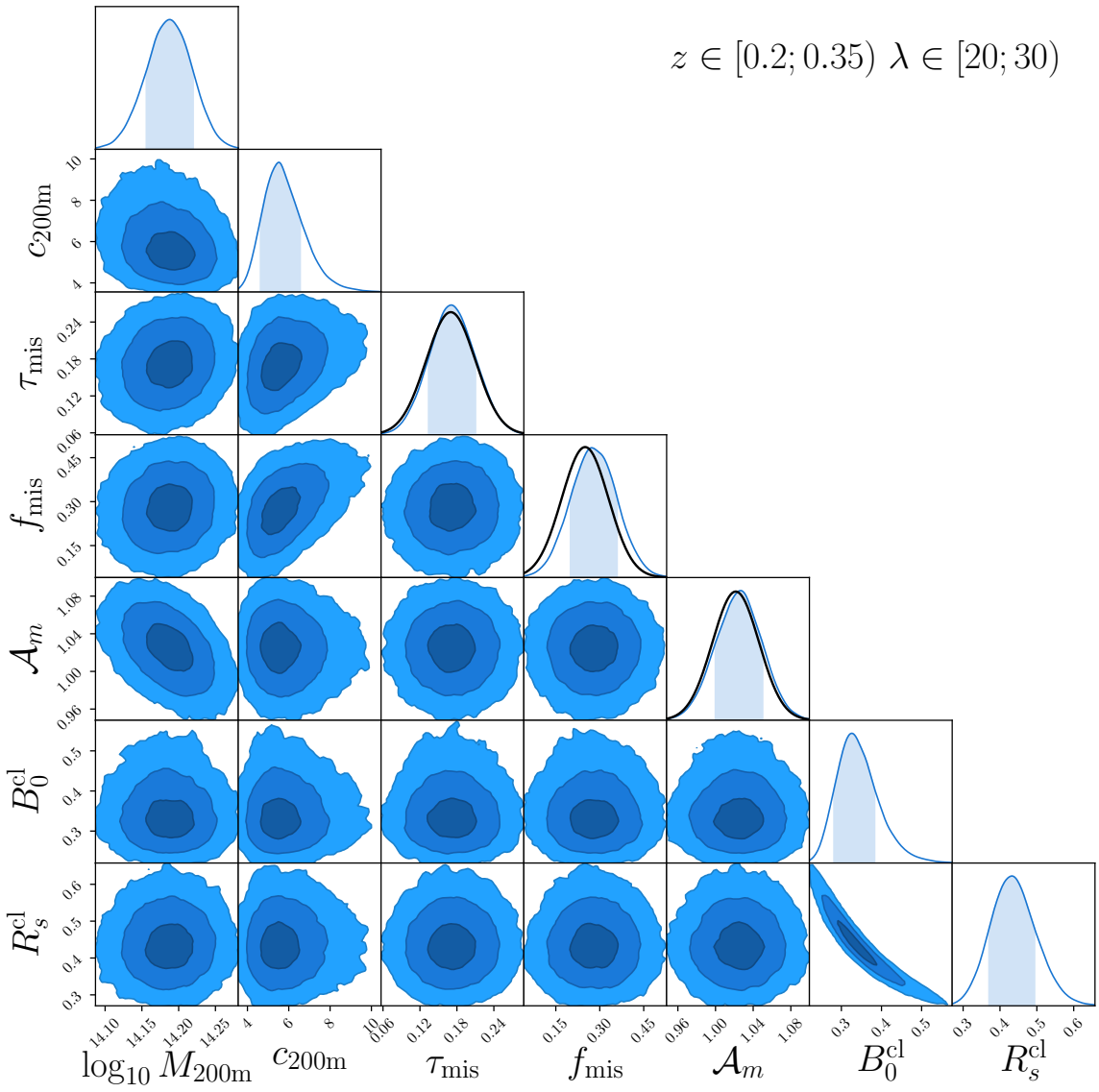

Figure 2.10: Posteriors for the parameters describing the lensing profile $\Delta\Sigma$ and the boost factor profile $\mathcal{B}$ for the bin $z \in [0.2, 0.35)$, $\lambda \in [20, 30)$. Contours show the 1σ, 2σ, and 3σ confidence areas. Black lines show the prior distributions. The mass presented here is uncalibrated, meaning it has not been corrected for modeling systematics, projection effects, or cluster triaxiality (see section 2.5.4).

Parameter	Description	Prior
$\log_{10} M_{200\mathrm{m}}$	Halo mass	$[11.0, 18.0]$
$c_{200\mathrm{m}}$	Concentration	$[0, 20]$
τ	Dimensionless miscentering offset	0.17 ± 0.04
f_{mis}	Miscentered fraction	0.25 ± 0.08
A_m	Shape & photo-z bias	Equation 2.45
B_0^{cl}	Boost magnitude	$[0, \infty]$
R_s^{cl}	Boost factor scale radius	$[0, \infty]$

Table 2.1: Parameters entering $\mathcal{L}(\Delta\Sigma)$ (Equation 2.48) and $\mathcal{L}(\mathcal{B})$(Equation 2.49) Flat priors are specified with limits in square brackets, Gaussian priors with means $\pm$ standard deviations.

The calibration correction described in 2.5.4 was applied to the recorded chains for each cluster subset. Specifically, for each point in the chain, we randomly sample the mass calibration factor $\mathcal{C}(\lambda, z)$ from its posteriors to adjust the mass. Further, we also apply the effect of triaxiality and projection effects (section 2.5.4.2), both of which add 2 per cent to the uncertainty of each mass. In practice this is written as

$$M = \mathcal{C}(\lambda, z)\mathcal{G}(0.96, 0.02)\mathcal{G}(1.02, 0.02) \times M_0\,, \tag{2.51}$$

where $\mathcal{G}$ is a Gaussian and M_0 are the "uncalibrated" masses in the chains. In this fashion, our final posteriors are properly marginalized over the uncertainty in the calibration factor $\mathcal{C}$ as well as triaxiality and projection effects.

In order to characterize the contribution of both statistical and systematic uncertainties to our final results we perform our analysis three different times with three different sets of assumptions. These three analyses we run are:

- `Full`: All parameters (concentration, Shear+photo-z, boost factors, miscentering) are allowed to vary within their priors. This is our fiducial analysis.
- `FixedAm`: $\mathcal{A}_m$ is set to the center of its prior distribution but all other parameters are allowed to vary. This determines the contribution from the shape and photo-z uncertainties.
- `OnlyMc`: Only mass and concentration are free. All other parameter priors are set to δ-functions at their central values. This represents our statistical uncertainty on the mass.

Table 2.2 contains the results of the `Full` analysis. Full posteriors from the cluster subset $z \in [0.2, 0.35)$ and $\lambda \in [20, 30)$ are shown in Figure 2.10. The corresponding data and best-fit model are shown in Figure 2.11, where we also demonstrate the combined effects of miscentering, boost factors, reduced shear, and multiplicative bias. The best fit model for each richness and redshift bin is over-plotted on top of the weak lensing data in Figure 2.4.

From the `Full` analysis we can see the contribution of the various systematics to our final results. The boost factors amount to a correction of $\approx$ 2 per cent to $\Delta\Sigma$ at R = 1 Mpc. The posteriors on the miscentering parameters are equal to the priors, demonstrating that these parameters are only weakly constrained by the weak lensing data. In our earlier analysis (Melchior

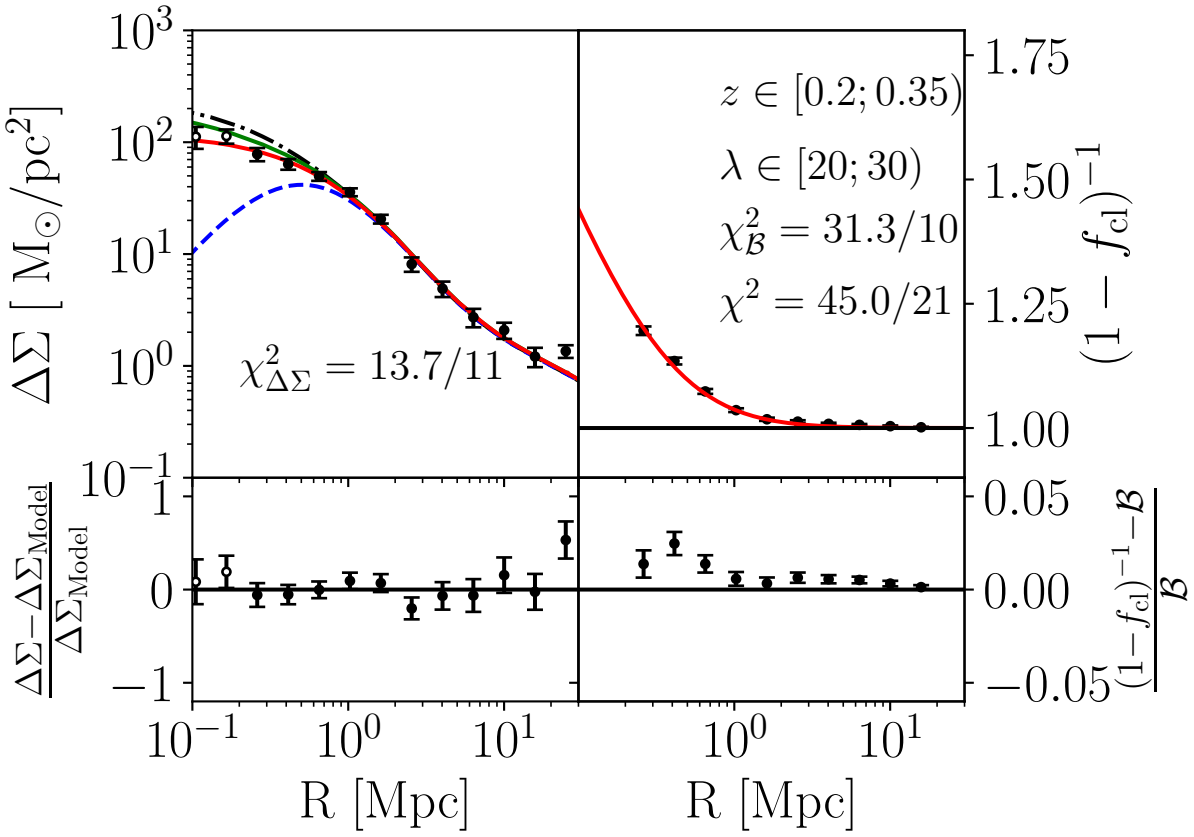

Figure 2.11: Fit with all components of the $\Delta\Sigma$ and $\mathcal{B}$ models for the cluster subset $z \in [0.2, 0.35)$ and $\lambda \in [20, 30)$. The top two panels show the best fit models in red compared to the data. Unfilled points are not included in the fit. *Top left*: the black dot-dashed line is $\Delta\Sigma_{\rm cen}$ while the blue dashed line is $\Delta\Sigma_{\rm mis}$. The weighted mean of these two yields the green solid line, and then applying the boost factor model, reduced shear, and multiplicative bias yields the final model in red. *Top right*: the red line is our NFW model for the boost factors. *Bottom*: the fractional difference between our data and models. The total χ^2 is 45 with 21 degrees of freedom, which is acceptable despite the imperfect fit of our simple model to the boost factors. The boost factors are measured from the data with small uncertainty, which is why the small mismatch with respect to the best-fit model causes a relatively large χ^2 but negligible effect on the recovered mass.

λ	$z \in [0.2, 0.35)$	$z \in [0.35, 0.5)$	$z \in [0.5, 0.65)$
$[20, 30)$	$14.191 \pm 0.013 \pm 0.032$	$14.162 \pm 0.013 \pm 0.033$	$14.083 \pm 0.015 \pm 0.048$
$[30, 45)$	$14.477 \pm 0.014 \pm 0.031$	$14.446 \pm 0.014 \pm 0.031$	$14.456 \pm 0.015 \pm 0.041$
$[45, 60)$	$14.608 \pm 0.011 \pm 0.044$	$14.643 \pm 0.011 \pm 0.044$	$14.648 \pm 0.016 \pm 0.056$
$[60, \infty)$	$14.913 \pm 0.014 \pm 0.038$	$14.899 \pm 0.015 \pm 0.038$	$14.879 \pm 0.023 \pm 0.061$

Table 2.2: Calibrated masses for each richness-redshift stack. All masses are in units of $\log_{10}\mathrm{M}_\odot$ using the $M_{200\mathrm{m}}$ definition. Listed uncertainties are split into the symmetric 68 per cent confidence intervals of the systematic and statistical components, in that order. Adding the two in quadrature gives the total uncertainty on the mass.

et al., 2017) we found a weak correlation between f_{mis} and M, which did not appear in this work. This was due to our use of the Diemer & Kravtsov (2014) M–c relation. We determined this by running one additional configuration in which the concentration was fixed by the Diemer & Kravtsov (2014) M–c relation, thus increasing the correlation between f_{mis} and M. At present, the contribution of miscentering to the mass is sub-dominant to other sources of systematic uncertainties in our final error budget (cf. Table 2.6). The multiplicative bias $\mathcal{A}_m$ follows the prior and is degenerate with mass.

The `OnlyMc` likelihood evaluation allows us to quantify the statistical and systematic uncertainties of the fiducial analysis. The difference in quadrature between the uncertainties in the `Full` and `OnlyMc` configurations represents the total systematic contribution to the error budget, while the `OnlyMc` alone provides the statistical contribution. The central values for each cluster subset along with statistical and systematic contributions to the uncertainties are presented in Table 2.2.

2.6 The mass–richness–redshift relation

The quantity we aim to constrain in this paper is the mean mass $\mathcal{M}(\lambda, z)$ of clusters of galaxies at a given observed richness λ and redshift z, similar to what was done in Melchior et al. (2015). Note that this is different from constraining the mean (and possibly distribution) of richness at given mass, or the full distribution of mass at given richness, as done in e.g. Simet et al. (2017b); Murata et al. (2017). In particular, we neither constrain nor require a model of the intrinsic scatter in richness, hence making this analysis largely independent from the choices in subsequent cluster cosmology studies based upon it.

We note that an assumed value of the intrinsic scatter is used in two places: the semi-analytic covariance matrices described in section 2.3.2.5 and the calibration described in section 2.5.4.1. For the covariance, this assumption had a negligible effect compared to the shape noise. While the overall calibration did depend on the amount of scatter, we took a conservative approach by treating the difference in calibration between assuming 15 per cent and 45 per cent scatter as a systematic uncertainty. In this way, our final results are not sensitive to the amount of assumed intrinsic scatter.

2.6.1 Modeling the mass–richness relation

We fit a redshift-dependent power-law relation between cluster richness and cluster mass. Specifically, we set

$$\mathcal{M}(\lambda, z) \equiv \langle M \,|\, \lambda, z \rangle = M_0 \left(\frac{\lambda}{\lambda_0}\right)^{F_\lambda} \left(\frac{1+z}{1+z_0}\right)^{G_z}, \tag{2.52}$$

where M_0, F_λ, and G_z are our model parameters. We select pivot values $\lambda_0 = 40$ and $z_0 = 0.35$, which are very near the median values of the cluster sample. Note $\mathcal{M}$ is the expectation value of the mass of a halo as a function of richness and redshift.

The expected mass of a given cluster subset k represents a weighted mean of the masses of the individual clusters in that stack. We then have

$$\mathcal{M}_k = \frac{\sum_{j\in k} \hat{w}_j \mathcal{M}(\lambda_j, z_j)}{\sum_{j\in k} \hat{w}_j}. \tag{2.53}$$

We take the weight $\hat{w}_j$ of the jth cluster to be the sum of the weights of all lens–source pairs $w_{j,i}$ around that cluster from 0.3 Mpc and above and verified that the choice of radial range does not affect our recovered masses. Individual cluster weights $\hat{w}_j$ differ from unity. This is because 1) the lensing weight of each lens–source pair given by Equation 2.11 depends on the cluster redshift, and 2) in a given radial bin there are more sources associated with low redshift clusters since that bin subtends a larger angle on the sky compared to the same bin around a high redshift cluster. In other words, the mass in the bin is skewed toward the average mass of the lower redshift clusters in the bin.

The impact of the weak lensing weights on the stacked mass estimates can be characterized by the ratio

$$a = \frac{\mathcal{M}_0}{\mathcal{M}_{\hat{w}}}. \tag{2.54}$$

We report the quantity $\log a$ in Table 2.3. We chose to report $\log a$ rather than a which has the computational advantage that one need only add $\log a$ to the mass values in Table 2.2 to arrive at an estimate of the mean cluster masses of cluster in a bin with unit weighting (as opposed to mean weak lensing weighted masses). The corrections in Table 2.3 are used to correct the recovered cluster masses to unit-weighted masses in the DES Y1 analysis of cluster abundances (DES collaboration, in preparation).

2.6.2 Mass covariance

The purpose of our different chain configurations (`Full`, `FixedAm`, and `OnlyMc`) is to allow us to estimate the contribution of each systematic to the final uncertainty on the mass calibration parameters M_0, F_λ, and G_z. In our analysis there are seven sources of systematic uncertainty: multiplicative shear bias, multiplicative photo-z bias, miscentering, boost factors, modeling systematics, triaxiality and projection.

We discuss how we combine all systematics to arrive at a full covariance matrix between our bins that respects the covariance in our systematic error budget. The reader will recall that the

λ	$z \in [0.2, 0.35)$	$z \in [0.35, 0.5)$	$z \in [0.5, 0.65)$
$[20, 30)$	-1.372×10^{-3}	-8.744×10^{-4}	-4.501×10^{-4}
$[30, 45)$	-2.979×10^{-3}	-3.278×10^{-3}	-6.660×10^{-4}
$[45, 60)$	-8.258×10^{-4}	-7.856×10^{-5}	-1.903×10^{-3}
$[60, \infty)$	3.043×10^{-3}	-4.061×10^{-3}	6.264×10^{-3}

Table 2.3: The logarithm of the mean mass correction factor $\log_{10} a$ from Equation 2.54. This represents a correction to the stacked cluster masses due to the fact that different clusters contribute to the measured mass in a different way than they contribute to $\Delta\Sigma$.

original chains we obtain from fitting the weak lensing profiles are processed via Equation 2.51 to account for the effect of calibration, triaxiality, and projection effects. If we let M_0 denote our unprocessed chains, and M denote the chains post-processing, in order to derive the statistical uncertainty only in our mass measurement we generate a new chain $\tilde{M}$ via

$$\tilde{M} = \bar{C} \times 0.96 \times 1.02 \times M_0 \,. \tag{2.55}$$

The difference in the variance between chain M in Equation 2.51 and that of chain $\tilde{M}$ represents the uncertainty associated with calibration, triaxiality, and projection effects. We will use the M without a $\sim$ to denote the chains post-processed as per Equation 2.51, and $\tilde{M}$ for chains post-processed as above.

The `OnlyMc` chain configuration contains only the statistical uncertainty in our analysis. For this reason, the covariance matrix describing the masses in this configuration is diagonal. We define the statistical uncertainty of the ith mass $\sigma^2_{i,\mathrm{stat}}$

$$\sigma^2_{i,\mathrm{stat}} = \mathrm{Var}\left(\tilde{M}_i^{\mathrm{OnlyMc}}\right) \,. \tag{2.56}$$

We also isolate the uncertainty associated with shear and photo-z systematics. To do so, we note the `Full` chain configuration includes all sources of uncertainty. Consequently, the difference in the variance between this chain and the `FixedAm` chain represents the uncertainty associated with shear and photo-z systematics. Therefore, we define

$$\sigma^2_{i,\mathrm{S+Pz}} = \mathrm{Var}(M_i^{\mathrm{Full}}) - \mathrm{Var}(M_i^{\mathrm{FixedAm}}) \,. \tag{2.57}$$

Finally, the uncertainty associated with modeling systematics (calibration, triaxiality, and projection effects) takes the form

$$\sigma^2_{i,\mathrm{mod}} = \mathrm{Var}(M_i^{\mathrm{Full}}) - \mathrm{Var}(\tilde{M}_i^{\mathrm{Full}}) \,. \tag{2.58}$$

By construction, the full uncertainty $\sigma^2_{i,\mathrm{Tot}}$ satisfies

$$\mathrm{Var}(M_i^{\mathrm{Full}}) = \sigma^2_{i,\mathrm{stat}} + \sigma^2_{i,\mathrm{S+Pz}} + \sigma^2_{i,\mathrm{mod}} \,. \tag{2.59}$$

We now define three different covariance matrices. First, $\mathsf{C}^{\rm stat}$ is diagonal, with $\mathsf{C}^{\rm stat}_{ii} = \sigma^2_{i,{\rm stat}}$. When we fit the weak lensing masses using this covariance matrix, we recover the statistical uncertainty in our scaling relation parameters. Second, $\mathsf{C}^{\rm S+Pz}$ is defined via

$$\mathsf{C}^{\rm S+Pz}_{ii} = \sigma^2_{i,{\rm stat}} + \sigma^2_{i,{\rm S+Pz}} \tag{2.60}$$

$$\mathsf{C}^{\rm S+Pz}_{ij} = \left[\sigma^2_{i,{\rm S+Pz}}\sigma^2_{j,{\rm S+Pz}}\right]^{1/2} . \tag{2.61}$$

Note the shear and photo-z components of the uncertainty are perfectly correlated across all bins. Fitting the weak lensing mass with this covariance matrix, and subtracting the statistical uncertainty in quadrature, enables us to calculate the uncertainty in our scaling relation parameters associated with shear and photo-z systematics. Third, $\mathsf{C}^{\rm Full}$ is defined via

$$\mathsf{C}^{\rm Full}_{ii} = \sigma^2_{i,{\rm stat}} + \sigma^2_{i,{\rm S+Pz}} + \sigma^2_{i,{\rm mod}} \tag{2.62}$$

$$\mathsf{C}^{\rm Full}_{ij} = \left[\left(\sigma^2_{i,{\rm S+Pz}} + \sigma^2_{i,{\rm mod}}\right)\left(\sigma^2_{j,{\rm S+Pz}} + \sigma^2_{j,{\rm mod}}\right)\right]^{1/2} . \tag{2.63}$$

Just like the shear and photo-z calibration uncertainties, modeling systematics are assumed to be perfectly correlated across all bins. The posteriors for the scaling relation parameters derived with this covariance matrix represent our full error budget, and the difference in quadrature between these errors and those obtained using the covariance matrix $\mathsf{C}^{\rm S+Pz}$ give us the error budget associated with modeling systematics.

2.6.3 Likelihood for the mass–observable relation

We model the likelihood of the recovered weak lensing masses as Gaussian in the log. This is illustrated in Figure 2.12, which shows the posterior for each of the 12 cluster bins with $\lambda \geq 20$, along with the corresponding Gaussian approximation. We rely on the $\lambda \geq 20$ cluster bins exclusively as it is only these systems for which we are confident we can unambiguously map halos to clusters and vice-versa. The full likelihood function is

$$\ln \mathcal{L}(M_{\rm obs}\,|\,M_0, F_\lambda, G_z) \propto -\frac{1}{2}(\Delta \log M)^T\ \mathsf{C}_M^{-1}\ (\Delta \log M)\,, \tag{2.64}$$

where C_M is the covariance between the mass bins for a particular configuration (see section 2.6.2). In the above equation $\Delta \log M$ is the difference between the measured mass of each cluster subset and our model for the expected mass given as per Equation 2.52. Thus, for the kth bin

$$\Delta \log M_k = \log M_k - \log \mathcal{M}_k\,. \tag{2.65}$$

We sample the posterior of the MOR parameters using *emcee* with 48 walkers taking 10000 steps each, discarding the first 1000 steps of each walker as burn-in. Table 2.4 summarizes the posteriors of our model parameters, while Figure 2.13 shows the corresponding confidence contours. All parameters in the M–λ–z relation have flat priors.

We explicitly enforce correlated uncertainties of shear, photo-z, modeling systematics, and triaxiality and projection effects. Miscentering and boost factors are considered independent

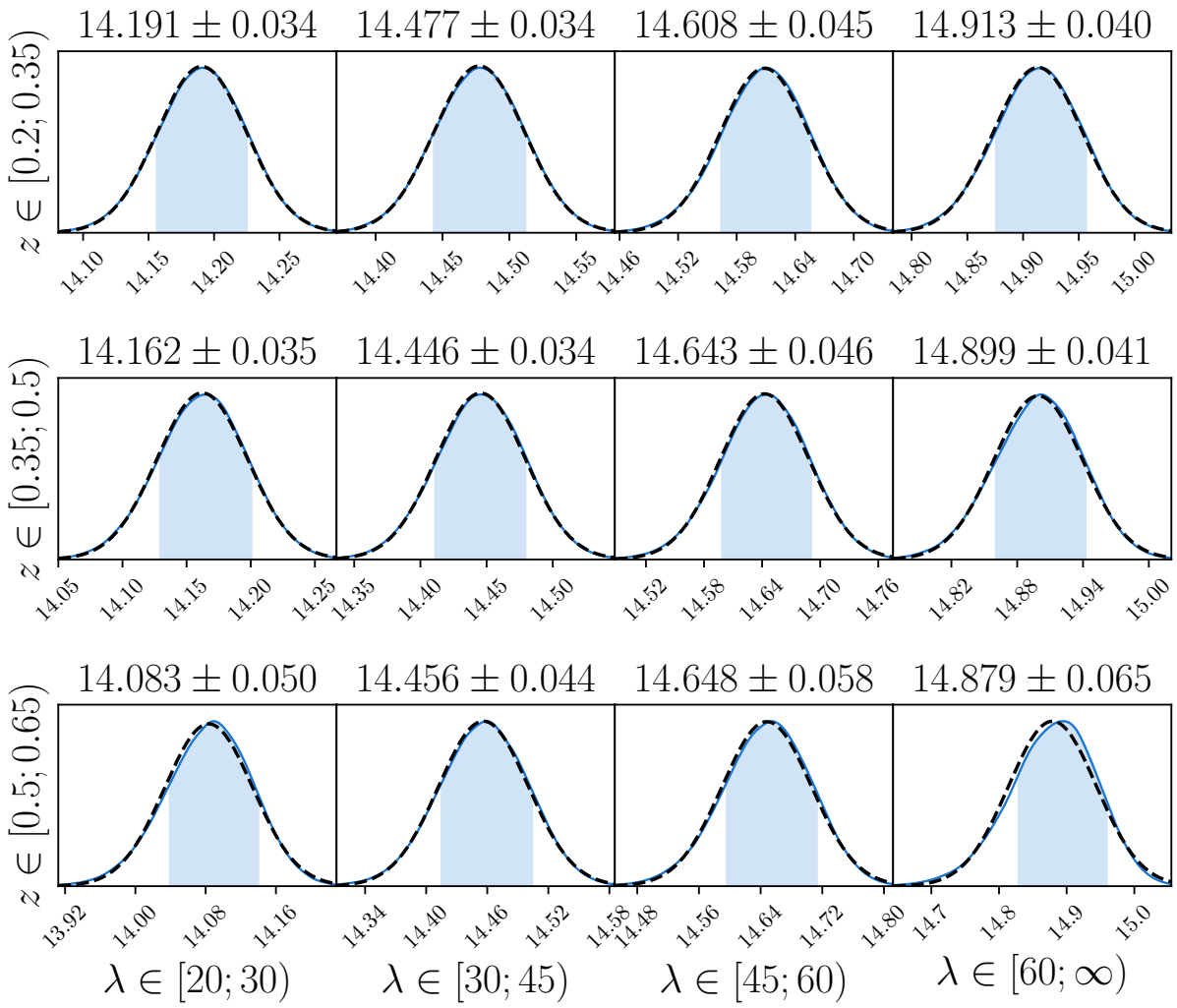

Figure 2.12: The calibrated posteriors of the masses for each cluster stack. Uncertainties appear above each panel, and are highlighted by the blue shaded regions. Gaussian approximations to these posteriors appear as black dashed lines.

Parameter	Description	Posterior
$\log_{10} M_0$	Mass pivot	$14.489 \pm 0.011 \pm 0.019$
F_λ	Richness scaling	$1.356 \pm 0.051 \pm 0.008$
G_z	Redshift scaling	$-0.30 \pm 0.30 \pm 0.06$

Table 2.4: Parameters of the M–λ–z relation from Equation 2.64 with their posteriors. The mass is defined as $M_{200\mathrm{m}}$ in units of $\mathrm{M}_\odot$. The pivot richness and pivot redshift correspond to the median values of the cluster sample. Uncertainties are the 68 per cent confidence intervals and are split into statistical (first) and systematic (second).

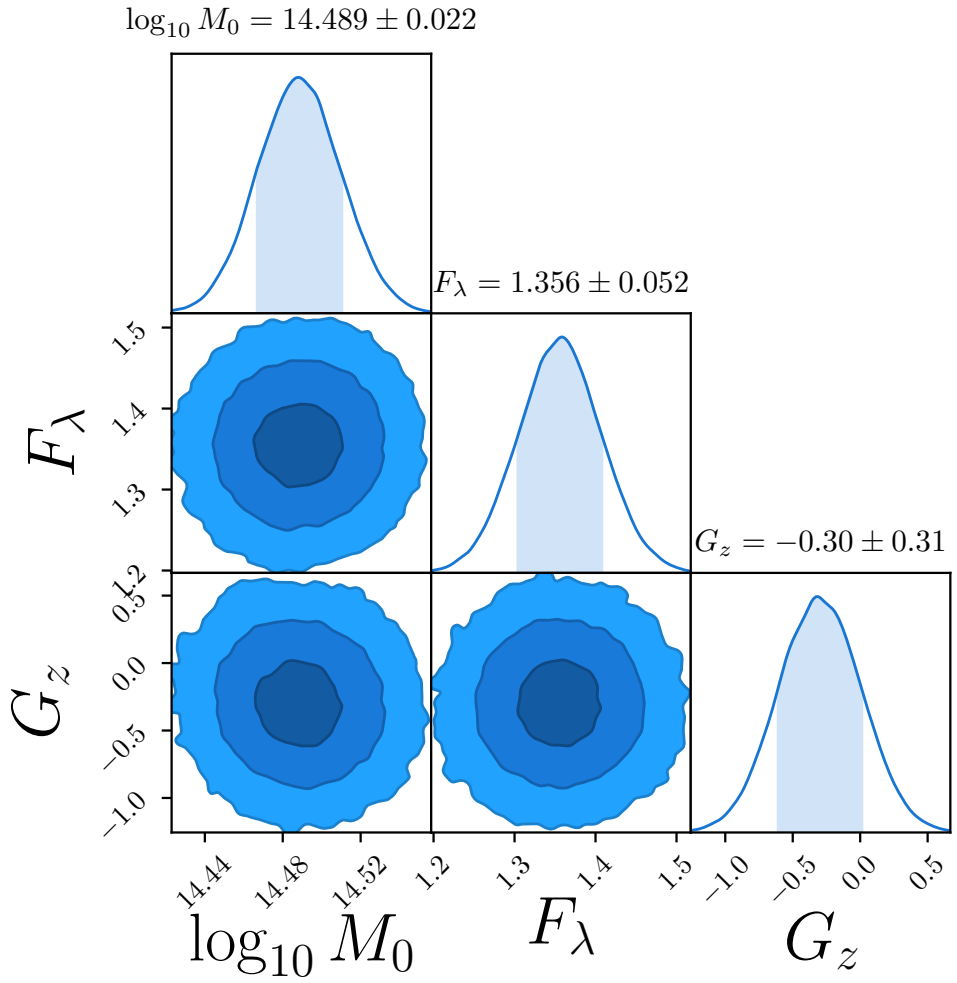

Figure 2.13: Parameters of the M–λ–z relation. Contours show the 1σ, 2σ and 3σ confidence areas from the `Full` run.

across cluster subsets. These independent uncorrelated systematics will tend to average out across bins.

In order to distinguish between the systematic and statistical contribution to the error budget on the M–λ–z relation parameters, we repeat the analysis using the statistical errors from the `OnlyMc` run. That is, we calculate Equation 2.64 using only the uncertainties measured from the `OnlyMc` run, or the $\mathsf{C}_M^{\rm stat}$ covariance matrix. The central values of the measured masses from the `OnlyMc` run are nearly identical to the `Full` run, as are the parameters in the M–λ–z relation. The difference in quadrature between the two uncertainties represents the systematic contribution while the excess uncertainty from the `OnlyMc` run is the statistical contribution. These uncertainties are reported in Table 2.4.

Our results imply that galaxy clusters of richness $\lambda = 40$ at redshift $z = 0.35$ have a mean mass of $\log_{10}\mathcal{M} = 14.489 \pm 0.011$ (stat) ± 0.019 (sys). The richness scaling is slightly steeper than linear at $F_\lambda = 1.356 \pm 0.051$ (stat) ± 0.008 (sys), while the mass shows a weak redshift dependence of $G_z = -0.30 \pm 0.30$ (stat) ± 0.06 (sys) consistent with no evolution. This amounts to a 5.0 per cent calibration (2.4 per cent statistical, 4.3 per cent systematic), of the M–λ–z relation.

In Melchior et al. (2017), we found that the dominant systematic uncertainty stemmed from shear and photo-z systematics, as was the case in Simet et al. (2017b). By repeating our analysis with the `FixedAm` run, which includes all systematics *except* $\mathcal{A}_m$, we are able to quantify the contribution from these sources. We found that the posterior distributions from the M–λ–z relation are significantly reduced, and that shear and photo-z systematics alone account for 48 per cent of the systematic uncertainty. This means that the remaining 52 per cent of the systematic uncertainty is due to modeling systematics, projection effects, and cluster triaxiality.

2.6.3.1 Alternative model using $\xi_{\rm lin}$

Hayashi & White (2008) used a similar model to ours, but with the *linear* matter correlation function for their 2-halo term. This causes very different behavior near the 1-halo to 2-halo transition region, which can affect the fitting procedure, as discussed in Melchior et al. (2017). We repeated our entire analysis, including recomputing the calibration, using $\xi_{\rm lin}$ in place of $\xi_{\rm nl}$. The masses of the stacks changed by less than 1 per cent, as did the normalization of the M–λ relation $\log_{10} M_0$. This means that our approach of calibrating the masses is largely robust to our choice of model.

2.6.3.2 Additional tests

We performed additional tests to verify our results. To ensure against possible small-scale systematic effects, we repeated our analysis with a more conservative radial cut of 500 kpc rather than 200 kpc. The resulting M–λ–z relation changed only in the mass scale, with M_0 changing by 0.2σ.

We also tested against possible differences in modeling systematics between large and small scales. By dividing each $\Delta\Sigma$ profile at 2 Mpc into large and small scale samples we could fit these regimes independently. While the constraining power was greatly diminished, the recovered masses were consistent with each other and the fiducial value within errors. No trend was

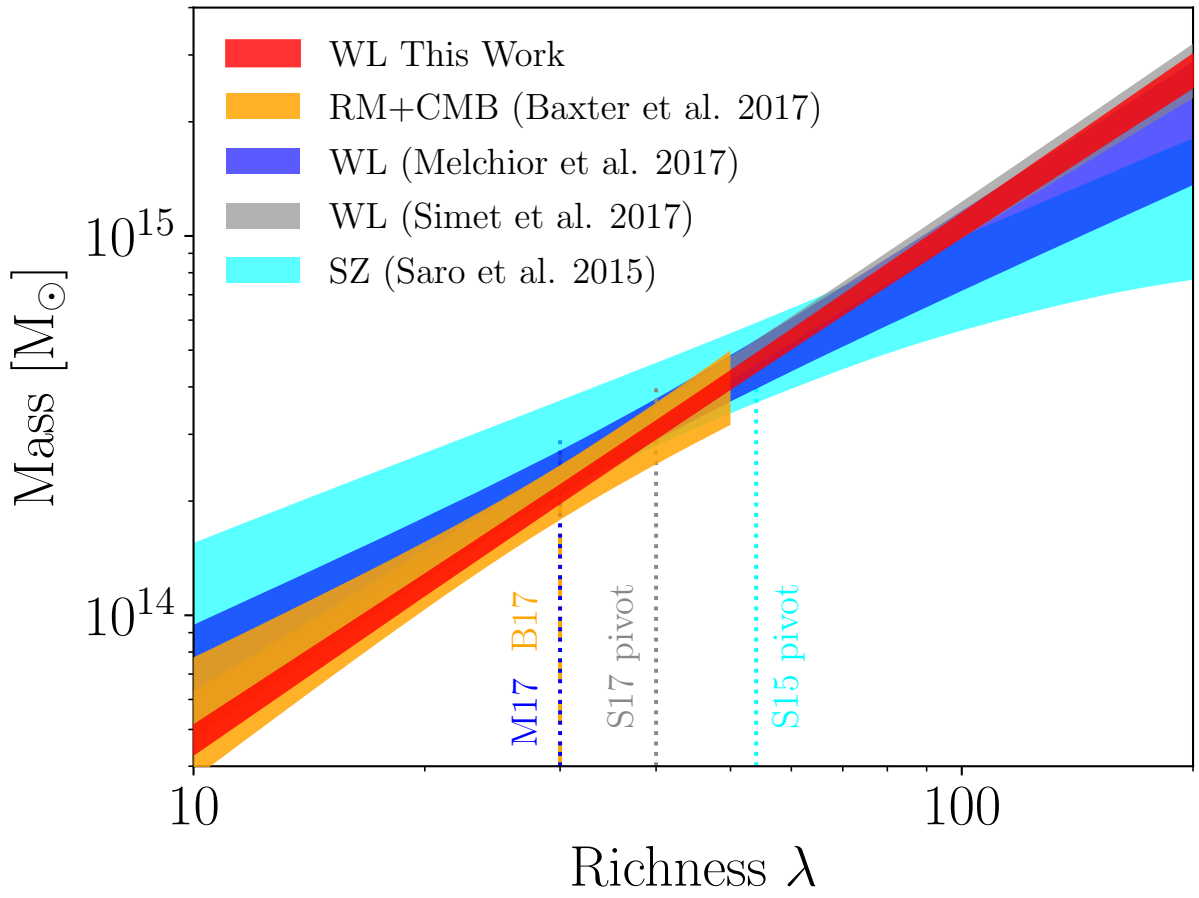

Figure 2.14: Best fit model for M–λ relation evaluated at the pivot redshift of our model, z_0 = 0.35, compared to other measurements. Our pivot richness is at λ_0 = 40. The previous DES result is in blue, from Melchior et al. (2017), while the relation measured in this analysis is in red. The analysis by Baxter et al. (2018) in orange used the same clusters as this work and found a consistent scaling relation over the richness range it probed.

observed in the differences between the recovered masses in any of these tests compared to the fiducial masses in Table 2.2.

Lastly, we tested an extension of Equation 2.52 where $F_\lambda(z) = F_{\lambda,0} + zF_{\lambda,1}$ and found $F_{\lambda,1}$ consistent with 0 at the 1.2σ level. Therefore, if any redshift evolution exists in the richness scaling, we are unable to resolve the behavior at present.

2.7 Comparison to Results in the Literature

We compare our calibration of the M–λ relation to previous results from the literature. The specific richness–mass relations we consider are summarized in Table 2.5, and we describe below the origin of each of these.

- Melchior et al. (2017) was the precursor to this analysis. In that work, we calibrated the mass–richness relation of redMaPPer clusters in the DES Science Verification data. A detailed description of the changes between that analysis and this one appears in the next section.

- Baxter et al. (2018) used the lensing of the Cosmic Microwave Background as measured by the South Pole Telescope to measure the mass–richness relation of DES Y1 redMaPPer clusters. Their analysis focused on 7066 clusters with richness $20 \leq \lambda \leq 40$. The upper limit was set to avoid potential biases in the recovered masses from contamination by thermal Sunyuaev-Zel'dovich emission by the clusters.

- Simet et al. (2017b) measured the mass–richness relation of redMaPPer clusters found in the Sloan Digital Sky Survey (SDSS). While their analysis is similar in spirit to ours, there are numerous methodological differences, including modeling choices (Simet et al. only fit the 1-halo term in the lensing profile), different radial scales used in the fit, a different shape catalog, and different photometric redshift catalogs.

- Murata et al. (2017) measured the richness–mass relation of SDSS redMaPPer clusters assuming a *Planck* cosmology. We compute the mean mass at $\lambda = 40$ as well as the local slope at this point in the scaling relation.. As demonstrated in Murata et al. (2017), their work and Simet et al. (2017b) are consistent with each other, despite the fact that they used different models for $\Delta\Sigma$, different radial scales and slightly different richness bins. Of special note is the fact that while Simet et al. (2017b) modeled only the 1-halo term using an NFW profile (along with a calibration step to correct for any biases introduced by this choice), Murata et al. (2017) used an emulator approach to simultaneously model the 1-halo and 2-halo terms of the lensing profile. In addition, the authors used both lensing profiles and cluster abundances, and assumed a fixed mass–concentration relation. These significant differences add information to the mass–richness relation, allowing for more precision than using lensing alone, as is the case in our analysis.

- Baxter et al. (2016) analyzed the cluster clustering of SDSS redMaPPer clusters. By measuring the angular correlation function of clusters they were able to constrain the amplitude

Authors	Description	$\log\langle M\|\lambda = 40, z = 0.35\rangle$ [$M_\odot$]	Richness scaling index F_λ
This work	weak lensing calibration using DES Y1	14.489 ± 0.022	1.356 ± 0.052
Melchior et al. (2017)	weak lensing calibration using DES SV	14.540 ± 0.067	1.12 ± 0.21
Baxter et al. (2018)	CMB lensing calibration using DES Y1	14.49 ± 0.31	1.24 ± 0.30
Baxter et al. (2016)	cluster clustering using SDSS	14.7 ± 0.1	1.18 ± 0.16
Simet et al. (2017b)	weak lensing calibration using SDSS	14.48 ± 0.03	1.30 ± 0.09
Murata et al. (2017)	weak lensing calibration using SDSS	14.543 ± 0.013	1.17 ± 0.07
Farahi et al. (2016)	pairwise velocity dispersion using SDSS	14.42 ± 0.10	1.31 ± 0.14
Saro et al. (2015)	SZE mass calibration using SPT and DES SV	14.44 ± 0.05	0.91 ± 0.18
Mantz et al. (2016)	weak lensing of individual WtG clusters	14.42 ± 0.11	1.36 ± 0.21
Geach & Peacock (2017)	CMB lensing calibration using SDSS	14.37 ± 0.05	0.74 ± 0.3

Table 2.5: redMaPPer scaling relation comparisons from the literature. Of note, the Simet et al. (2017b) results have changed slightly (Simet, private communication). We evaluate $\log_{10}\langle M|\lambda = 40, z = 0.35\rangle$ of the other scaling relations in order to compare them to our result, applying the richness correction given by Equation 2.66. When necessary, we use the method presented in Evrard et al. (2014) to convert from richness–mass relations to mass–richness relations. All masses are $M_{200\mathrm{m}}$.

of the mass scaling relation to 18 per cent, in which their dominant systematic was uncertainty in the bias–mass relation.

- Farahi et al. (2016) measured masses using stacked pairwise velocity dispersion measurements of SDSS redMaPPer clusters. Their measurements serve as a good cross check against other analyses of SDSS clusters, but found that they are ultimately less precise due to large uncertainties in velocity bias.

- Saro et al. (2015) measured the mass–richness relation of galaxy clusters by assuming a *Planck* cosmology to determine the observable–mass relation of clusters from the South Pole Telescope (Bleem et al., 2015). They then matched these SPT clusters to redMaPPer clusters from the DES Science Verification data, and use the overlap sample to determine the richness–mass relation. We invert the relation using the method of Evrard et al. (2014) in order to show the comparison in Figure 2.14.

- Mantz et al. (2016) compared the scaling relation measured from the Weighting the Giants mass estimates for individual redMaPPer clusters in SDSS from Applegate et al. (2014b) to that of the Simet et al. (2017b) analysis. They found the two scaling relations in good agreement, which is also the case when compared to our measurement.

- Geach & Peacock (2017) constrained the mass–richness relation of redMaPPer clusters found in SDSS using convergence profiles measured from *Planck* data. They constrain the normalization of the scaling relation at the $\sim 11.5\%$ level, but are unable to replicate this for the scaling index. Of course, this measurement will improve in the future as both optical cluster catalogs expand and CMB lensing maps improve. Their method will provide a necessary cross check on any potential mass bias due to unknown systematics.

Table 2.5 summarizes these scaling relations. Critically, the richness definition λ is sensitive to the details of image processing, source detection, choice of magnitudes, etc, and can therefore vary systematically from one survey to the next. We explicitly correct for this impact cross-matching DES Y1 clusters to DES SV and SDSS redMaPPer clusters, and measuring the richness offset. We find

$$\lambda_{\rm DES\ SV} = (1.08 \pm 0.16)\lambda_{\rm DES\ Y1} \tag{2.66}$$

$$\lambda_{\rm SDSS} = (0.93 \pm 0.14)\lambda_{\rm DES\ Y1}\,. \tag{2.67}$$

In these equations, the error is the standard deviation in the richness ratio, not the error on the mean. We apply these corrections to the SDSS and DES SV scaling relations before comparing to our result. So, for instance, if the scaling relation for data set X takes the form

$$\langle M|\lambda_X\rangle = A\lambda_X^\alpha \tag{2.68}$$

and the ratio $\lambda_X/\lambda_{\rm DES\ Y1} = r$, then the scaling relation for Y1 richnesses is

$$\langle M|\lambda_{\rm DES\ Y1}\rangle = Ar^\alpha\lambda_{\rm DES\ Y1}^\alpha. \tag{2.69}$$

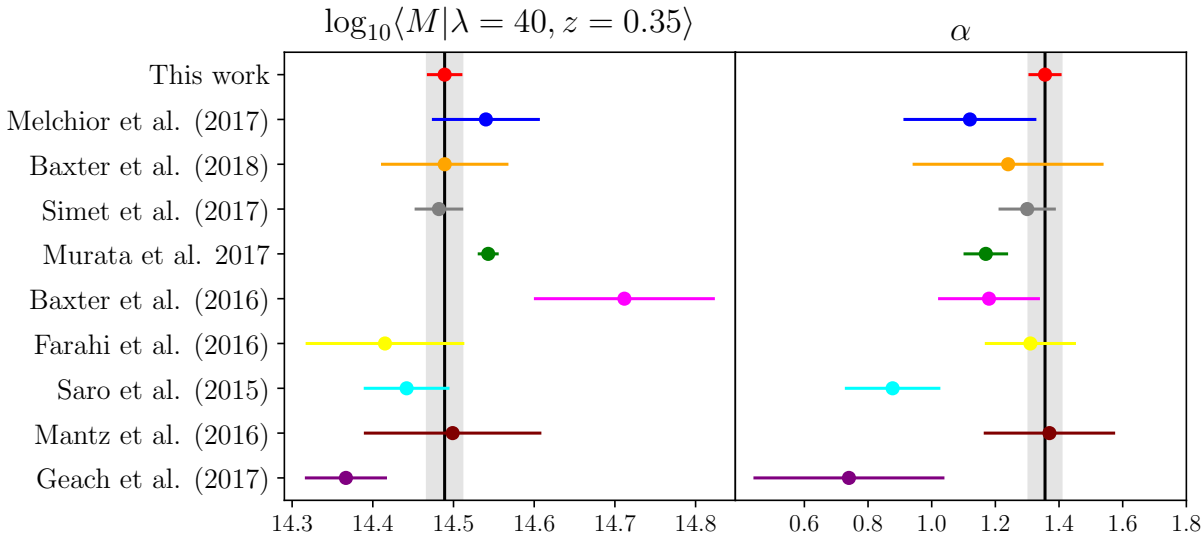

Figure 2.15: Comparison of the predicted mass at $\lambda = 40$ and $z = 0.35$ as well as the richness scaling relation between this work (*gray bands*) and other results from the literature.

Finally, all scaling relations that do not explicitly incorporate redshift evolution are transported from their quoted pivot redshift to our chosen pivot redshift $z = 0.35$ using our best fit redshift evolution.

Figure 2.15 and Table 2.5 show the mass at $\lambda = 40$ and $z = 0.35$ as well as the richness scaling index for each of the scaling relations described above.

Notably, we do not include some recent cluster calibration measurements. van Uitert et al. (2016) was not included because they used a different cluster finding algorithm and focused on significantly lower richness clusters. They were, however, able to constrain the mass–richness scaling relation at the 5% level using lensing and cluster-satellite correlations for low redshifts. Miyatake et al. (2018) and Medezinski et al. (2018a) calibrated the masses of SZ detected clusters in ACT and *Planck*, respectively. However, their primary focus was on investigating the hydrostatic mass bias, and not on cluster scaling relations, since their sample sizes were significantly smaller than ours. Nevertheless, we anticipate this narrative to change in the near future with with larger HSC cluster catalogs.

2.8 Systematic Improvements From DES SV to DES Year 1, and from Year 1 to Year 5

Our current analysis has multiple significant improvements relative to Melchior et al. (2017), the precursor to this work. Specifically:

- Shear calibration related errors on mass decreased from 4 per cent to 1.7 per cent, based primarily on the data driven correction of shear biases with METACALIBRATION. This

implies that the shear calibration uncertainty is no longer the dominant source of systematic error in our weak lensing analysis.

- The largest contribution to the systematic uncertainty is now photo-z errors. In the purely COSMOS-based calibration applied in this work, we find only a minimal improvement between SV and Y1.

- The $\approx$ 15 per cent increase in uncertainty due to using noisy jackknife estimates of the covariance matrix (Dodelson & Schneider, 2013) of the weak lensing profiles was entirely removed through the use of a semi-analytic covariance matrix.

- Uncertainty from modeling systematics decreased from 2 per cent to 0.73 per cent. In Melchior et al. (2017), the model calibration corrections were computed by stacking halos in mass bins. By contrast, our current analysis assigns richness according to a fiducial richness–mass relation, allowing us to stack in richness bins and to therefore accurately compute the correction for a richness bin. Relative to our SV analysis, the amplitude of this correction increased, while its uncertainty was reduced, from 1.00 ± 0.02 to 1.042 ± 0.004. The increase in the correction is primarily due to the richness binning of the halos as well as the fact that by using the Y1 covariance matrix, the impact of small differences on large scales between our analytical model and numerical simulations is amplified, leading to larger correction factors. More importantly, however, the uncertainty on this correction was greatly reduced. This is due to the semi-analytic covariance matrix as well as allowing for intrinsic scatter of the calibration factors. While we use the semi-analytic covariance matrix for calibration on the simulated profiles, the matrix does not adequately describe the uncertainty in any systematic differences between the model and real data. Additionally, we now fit for the associated systematic uncertainty from the dispersion in the calibration data. Both of these factors result in a decrease in the systematic error budget.

- The mass–concentration relation of galaxy clusters is allowed to float in this analysis, while it was held fixed in Melchior et al. (2017). The fixed mass–concentration relation in our SV analysis was necessitated by the relatively low S/N of the weak lensing measurements. By contrast, our current analysis enables us to marginalize over concentration, which in turn should make our results significantly more robust to the impact of baryonic physics in the cores of galaxy clusters.

All in all, the reduced statistical and systematic uncertainty in our analysis has reduced the error in the amplitude of the mass–richness relation from 11.2 per cent to 5.0 per cent. Unlike our analysis in Melchior et al. (2017), our current constraints are close to systematics limited. Without improved systematics between now and the end of the survey, the improved statistics of the Year 5 data will only decrease our total error budget from 5.0 per cent to $\approx$ 4 per cent. Evidently, further reducing systematic uncertainties in future weak lensing mass calibration analysis is imperative.

Photometric redshift errors currently dominate the systematic error budget. Significant improvements in the weak lensing mass calibration of galaxy clusters will require new algorithms

for calibrating photometric redshifts. A joint constraint from high-fidelity photo-zs of matched reference samples and clustering redshifts (Gatti et al., 2018; Davis et al., 2018), as done in Hoyle et al. (2018), is not feasible for higher redshift sources due to the limited redshift range of available spectroscopic or redMaGiC catalogs. Alternatively, source selection criteria that take into consideration photometric redshift uncertainty could lead to a desirable trade off between statistical and systematic uncertainty. With some combination of these approaches, reducing the photometric redshift uncertainty by a factor of two for a year 5 analysis seems plausible.

Related to this, in our current analysis we have assumed that all systematics are perfectly correlated across all redshift and richness bins. This is likely *too* conservative. In particular, photometric redshift systematics are unlikely to be perfectly correlated across all redshift bins: the selection of the source population, and their relative signal contribution as a function of source redshift, differ as a function of lens redshift. Adequately characterizing the covariance in the systematic uncertainties associated with photometric redshift errors seems like a relatively simple way to significantly decrease our systematic error budget. For instance, if one were to assume that the photometric redshift systematics were entirely uncorrelated, the associated systematic would be reduced by a factor of $1/\sqrt{3}$, rendering photometric redshift errors subdominant. This is clearly unrealistic, but it does illustrate that characterizing the covariance in the systematics may lead to significant reductions in the total error budget.

Following photometric redshift uncertainties, three different effects come in at the ≈ 2 per cent level: shear systematics, triaxiality effects, and projection effects. Of these, shear systematics are the least problematic. We fully expect shear calibration uncertainties will continue to decrease over the coming years, and they will no longer be a major source of error for cluster mass calibration. By contrast, the current systematic error estimates for triaxiality and projection effects clearly demonstrate that there is a significant need for a detailed study of these on weak lensing mass profiles, such as in the recent work of Osato et al. (2018).

Additional, but less urgent, upgrades to our analysis are also possible. For instance, following Murata et al. (2017), an emulator based approach to modeling the halo-matter correlation function or $\Delta\Sigma$ directly can potentially greatly reduce the modeling calibration and its contribution to the uncertainty. Centering errors will also continue to decrease as the availability of multi-wavelength data continues to increase.

Finally, systematics that we have thus far ignored need to be better addressed. For instance, intrinsic alignment by cluster member galaxies even if its effect is very small (Sifón et al., 2015), which impact membership dilution estimates. Likewise, a study of the impact on baryonic physics on our weak lensing calibration methodology is necessary. While we expect these sources of error to be subdominant in our present study, quantifying the systematic error associated with these effects will be increasingly important in the future.

2.9 Summary and Conclusions

We measured the stacked weak lensing signal of redMaPPer cluster in the DES Y1 data. The clusters were divided into 21 subsets of richness and redshift. The mean mass of each cluster stack was estimated for those subsets with $\lambda \geq 20$ and $0.2 \leq z \leq 0.65$. Our model incorporated:

Source of systematic	SV Amplitude Error	Y1 Amplitude Error
Shear measurement	4%	1.7%
Photometric redshifts	3%	2.6%
Modeling systematics	2%	0.73%
Cluster triaxiality	2%	2.0%
Line-of-sight projections	2%	2.0%
Membership dilution + miscen.	$\leq$ 1%	0.78%
Total Systematics	6.1%	4.3%
Total Statistical	9.4%	2.4%
Total	11.2%	5.0%

Table 2.6: Systematic error budget on the amplitude of the mass–richness relation as measured with the DES Y1 data compared to the DES SV result of Melchior et al. (2017). The shear (section 2.4.1), photo-z (section 2.4.3), modeling systematics (section 2.5.4.1), triaxiality, and projection effects (section 2.5.4.2) are conservatively taken to be perfectly correlated between cluster stacks. Membership dilution (section 2.4.2) and miscentering (section 2.5.2) are assumed to be independent. Statistical and systematic errors are added in quadrature to arrive at the total error.

- Shear measurement systematics (section 2.4.1),
- Source photometric redshift uncertainties (section 2.4.3),
- Source sample dilution by cluster members (section 2.4.2, section 2.5.3.1),
- Cluster miscentering (section 2.5.2),
- Model calibration systematics (section 2.5.4),
- Triaxiality & projection effects (section 2.5.4.2).

The mean masses of the cluster subsets were used to determine the mean cluster mass as a function of richness and redshift according to Equation 2.52. We emphasize that the full analysis was performed blindly: the paper underwent internal review by the DES collaboration prior to unblinding, and no changes to the analysis were made post-unblinding.

We summarize our constraints on the scaling relation as follows: for clusters at our pivot richness of $\lambda_0 = 40$ and pivot redshift of $z_0 = 0.35$, the mean cluster mass is

$$M_0 = [3.081 \pm 0.075 \pm 0.133] \cdot 10^{14}\ \mathrm{M}_\odot. \tag{2.70}$$

The slope F_λ for the mass–richness relation is

$$F_\lambda = 1.356 \pm 0.051 \pm 0.008, \tag{2.71}$$

and the slope G_z governing the redshift evolution of the mass–richness relation is

$$G_z = -0.30 \pm 0.30 \pm 0.06, \tag{2.72}$$

where the first and second set of errors correspond to statistical and systematic errors, respectively. The full scaling relation is given by Equation 2.52. This scaling relation is in excellent agreement with, while being significantly more precise and accurate than, previous results from the literature: Mantz et al. (2016); Saro et al. (2015); Simet et al. (2017b); Baxter et al. (2018); Murata et al. (2017); Melchior et al. (2017).

The 5.0 per cent constraint on the amplitude of the mass–richness relation is systematics dominated, with our systematic error alone reaching 4.3 per cent. We stress the systematic uncertainty in the shear and photometric redshift catalogs have been extensively tested and validated, so we are confident our systematic error budget is robust. Halo triaxiality and line of sight projections are now an important contributor to the total systematic error, and represent a critical path for minimizing the overall error budget for future analyses

Mass calibration remains the limiting factor for the ability of testing cosmological models with cluster counts. Nevertheless, this work represents a significant step forward: we were able to reduce the systematic error budget from 6.1 per cent in DES Science Verification to 4.3 per cent in DES Year 1. While we will need to achieve similar level of improvements for future analyses including DES Year 5 and LSST Year 1 to significantly improve upon our results, we are confident that we will be able to rise to the challenge: the story of weak lensing mass calibration is one of ever decreasing systematic errors, a trend that to this day shows no signs of abating.

Chapter 3

Validation of Weak Lensing Cluster Member Contamination Estimates from P(z) Decomposition

Scientific context In the previous chapter 2 a crucial step of the analysis, the correction for cluster member galaxies contaminating the source galaxy catalog, also known *boost factors* was left as a side note in section 2.4.2 to be performed, validated and presented in detail in a separate study. This contamination effect, already illustrated on Figure 2.11, results in an order unity modulation of the intrinsic true lensing signal on small scales, therefore it needs to be correctly accounted and corrected for in a weak lensing measurement aiming at precise and accurate cluster masses.

In this chapter of original research we now embark on the task to test and validate the performance of an estimator for the cluster member contamination, which was first used in the cluster weak lensing analysis of Melchior et al. (2017) performed with the DES Science Verification dataset, and was motivated by the approach of Gruen et al. (2014) and Dietrich et al. (2019). This method makes use of the the photometric redshift probability distribution function $P(z)$ of source galaxies, which it decouples into two non-negative terms corresponding to a contaminating and a background component and estimates the contamination rate from the mixing factor between the two components. In the following we perform a full scale validation on a simulated sky survey approximately mirroring the observational properties of DES Y1, and apply the method to the actual DES Y1 data. We note that in addition to the current work of McClintock & Varga et al. (2019), the approach has since been used in a variety of cluster lensing studies in recent years by Chang et al. (2018b), Medezinski et al. (2018c), Stern et al. (2018), Pereira et al. (2020).

Authorship and publication The research presented in this chapter has been conducted within the Dark Energy Survey, and was published as *Dark Energy Survey Year 1 results: validation of*

weak lensing cluster member contamination estimates from P(z) decomposition in the Monthly Notices of the Royal Astronomical Society, volume 489, pages 2511–2524 (Varga et al., 2019).

The single leading author of this study is T. N. Varga (myself). According to DES publication policy the additional authors of the paper are divided into a first tier with direct contribution to the research in terms of data analysis or substantive discussion, second tier authors whose work indirectly enabled the analysis and whom are listed alphabetically, and a so-called *builder*-tier also listed alphabetically and which consits of people whose work contributed to DES infrastructure as a whole. In this study the first author was responsible for the full extent of leadership of the work, including data processing and the execution of the analysis. Further contributions from first tier authors are as follows: J. DeRose provided the *Buzzard* suite of mock sky-surveys constructed based on numerical simulations (DeRose et al., 2019), while T. McClintock contributed access to the Monte-Carlo Markov chains used in McClintock & Varga et al. (2019). We furthermore thank the advisory contribution of D. Gruen, E. Rozo, and S. Seitz.

All first tier authors contributed to the discussion section of the paper. In accordance with DES publication policy this paper has gone through internal review and collaboration wide review before the submitting to the journal.

The following part of this chapter reproduces the paper as it was accepted for publication in the journal with minor edits to suit the thesis format, while keeping the substantive part of the work unchanged. The acknowledgments of the original paper are reproduced in Appendix B.

Abstract Weak lensing source galaxy catalogs used in estimating the masses of galaxy clusters can be heavily contaminated by cluster members, prohibiting accurate mass calibration. In this study we test the performance of an estimator for the extent of cluster member contamination based on decomposing the photometric redshift $P(z)$ of source galaxies into contaminating and background components. We perform a full scale mock analysis on a simulated sky survey approximately mirroring the observational properties of the Dark Energy Survey Year One observations (DES Y1), and find excellent agreement between the true number profile of contaminating cluster member galaxies in the simulation and the estimated one. We further apply the method to estimate the cluster member contamination for the DES Y1 redMaPPer cluster mass calibration analysis, and compare the results to an alternative approach based on the angular correlation of weak lensing source galaxies. We find indications that the correlation based estimates are biased by the selection of the weak lensing sources in the cluster vicinity, which does not strongly impact the $P(z)$ decomposition method. Collectively, these benchmarks demonstrate the strength of the $P(z)$ decomposition method in alleviating membership contamination and enabling highly accurate cluster weak lensing studies without broad exclusion of source galaxies, thereby improving the total constraining power of cluster mass calibration via weak lensing.

3.1 Introduction

Galaxy clusters trace the highest peaks of the cosmic density field and their abundance and distribution constitutes a powerful cosmological probe (Allen et al., 2011; Dodelson et al., 2016). This mode of inference poses two major tasks: detecting galaxy clusters from observational data, and defining a mass–observable relation (MOR) to compare the observed cluster abundances with the predicted ones. The efficient pathways to identify galaxy clusters differ between the available wavelengths and targeted redshifts ranges: In optical wavelengths and low-redshifts ($z < 1$) clusters can be detected as an overdensity of quenched, red, early type galaxies (Koester et al., 2007; Rykoff et al., 2014), giving rise to the cluster mass proxy known as *optical richness* (λ). In other wavelengths clusters can be identified through the X-ray emission (Truemper, 1993; Mantz et al., 2010) or through the Sunyaev-Zeldovich effect (Sunyaev & Zeldovich, 1970, 1972; Bleem et al., 2015) induced by the presence of hot intra-cluster gas. While these methods are suitable to detect clusters, they do not provide a direct measure of their masses. The MOR must be calibrated using additional information.

The best method for calibrating cluster masses today is via weak gravitational lensing, as it is directly sensitive to the gravitational potential. For this reason weak lensing cluster mass calibration studies (von der Linden et al., 2014b,a; Applegate et al., 2014b; Hoekstra et al., 2015; Mantz et al., 2015; Okabe & Smith, 2016; Battaglia et al., 2016; Melchior et al., 2017; Simet et al., 2017b; Murata et al., 2017; Dietrich et al., 2019; McClintock & Varga et al., 2019) have become a necessary component of cluster cosmology analyses. Weak lensing mass estimates carry their own set of uncertainties, both systematic and statistical. It is expected that to fully utilize the statistical power of ongoing sky surveys, the amplitude of the MOR must be calibrated with at most a few percent total uncertainty (Weinberg et al., 2013). With the growing depth, area, and statistical power of various sky surveys the proper characterization of systematic un-

certainties is becoming the highest priority. Indeed, current analyses are systematics dominated (e.g. Table 6 of McClintock & Varga et al., 2019), meaning that to improve on the overall cosmological constraining power we have to improve our understanding of systematic errors in the mass calibration.

One important systematic impacting weak lensing analyses is the contamination of the source galaxy catalog by galaxies associated with the cluster. This contamination is a result of the uncertainty in photometric redshift estimates, as few-band surveys do not provide enough information to select a pure and close to complete background sample of galaxies. Contaminating galaxies dilute the measurement, requiring one to boost the raw signal to recover the true signal. Hence the effect is traditionally referred to as the *boost factor* (Sheldon et al., 2004; Applegate et al., 2014b; Hoekstra et al., 2015; Gruen et al., 2014; Simet et al., 2017b; Melchior et al., 2017; Medezinski et al., 2018d; Leauthaud et al., 2017). Conversely, when many-band photometric information is available, the contamination can be strongly reduced, but with increased observational cost (Applegate et al., 2014b).

Previous studies made use of multiple approaches in characterizing cluster member contamination: Sheldon et al. (2004) and Simet et al. (2017b) estimated the boost factor profiles from the transverse correlation of source galaxies around cluster centers, while Applegate et al. (2014b) and Medezinski et al. (2018d,b) utilized the color information in a "color-cut" method. Gruen et al. (2014) and Dietrich et al. (2019) estimated the correction factor from decomposing the source population into a cluster and background component. This latter method was expanded by Melchior et al. (2017) who estimated the contamination rate based on a decomposition of the photometric redshift probability distribution function ($P(z)$) estimates of source galaxies, which was also employed by Chang et al. (2018b) and Stern et al. (2018).

In this study we aim to validate the cluster member contamination estimates obtained through $P(z)$ decomposition, and provide a detailed description for the case of the DES Y1 cluster weak lensing analysis of McClintock & Varga et al. (2019). The structure of this paper is as follows: In section 3.2 we outline the framework and formalism of the $P(z)$ decomposition method, in section 3.3 we perform tests on simulated DES-like observations as well as actual DES data, and finally in section 3.4 we present the boost factor results used in the DES Y1 redMaPPer weak lensing cluster mass calibration (McClintock & Varga et al., 2019).

For the DES Y1 data analysis, we assume a flat ΛCDM cosmology with $\Omega_\mathrm{m} = 0.3$ and $H_0 = 70$ km s^{-1} Mpc^{-1}, with distances defined in physical coordinates, rather than comoving. The DES-like mock observations assume a flat ΛCDM cosmology with $\Omega_\mathrm{m} = 0.286$, $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_\mathrm{b} = 0.047$, $n_s = 0.96$, and $\sigma_8 = 0.82$.

3.2 P(z) decomposition formalism

Our aim is to estimate the cluster member contamination affecting weak lensing measurements. With an estimate of the contamination rate that has sufficiently low systematic and statistical uncertainty, we can correct for the bias in the raw weak lensing signal.

The present approach infers the fraction of contaminating cluster member galaxies f_cl from the photometric redshift $P(z)$ probability distribution function (p.d.f.) of the appropriately se-

lected and weighted source galaxies. By comparing the $P(z)$ of sources near clusters with the $P(z)$ of galaxies in field lines-of-sight we identify a feature indicative of the presence of cluster galaxies. The schematic of this approach is illustrated on Figure 3.1 for a selection of galaxy clusters. The three left panels show the average $P(z)$ of galaxies at three different transverse separations from the clusters, along with the reference $P(z)$ of field galaxies. The feature characteristic for the cluster member contamination is shown as a red Gaussian curve. We note that the position of the Gaussian is expected to be at a higher redshift than the redshift range of the targeted clusters, as only those cluster member galaxies contribute to the contamination whose redshift estimates are significantly biased high. The relative strength of this feature at different radii is taken as a tracer of the cluster member contamination rate profile $f_{\rm cl}(R)$, shown on the right panel of Figure 3.1.

3.2.1 Weak lensing formalism

Weak lensing analyses of galaxy clusters rely on a large sample of background *source* galaxies. The images of these background source galaxies are distorted due to the gravitational potential of the lens (the galaxy cluster), and thus can be used to trace the underlying matter distribution of the lens (Bartelmann & Schneider, 2001)

In most scenarios a catalog of background source galaxies is constructed from optical imaging data, and thus their exact distances are not known. To remedy this, photometric redshift algorithms are employed to provide an estimate of their redshifts. Such methods involve large uncertainty for individual galaxies, and potential bias for the ensemble, due to the limited information available (Hoyle et al., 2018; Hildebrandt et al., 2017). The uncertainty of photometric redshifts mean that the background source catalog can only be defined in approximate terms, such that it may also include foreground galaxies, and galaxies which are at the lens redshift. In contrast, the redshifts of galaxy clusters are typically known with very high precision either from spectroscopic follow up, or from the ensemble photometric redshift estimates of their red cluster member population (Rykoff et al., 2016:McClintock & Varga et al., 2019). In this analysis we neglect any uncertainty in the cluster redshift $z_{\rm lens}$, and consider that the uncertainty in the redshift of source galaxies $z_{\rm src}$ is captured in their $P(z_{\rm src})$ p.d.f.

The photometric redshift $P(z)$ assigned to a galaxy is determined by its observed properties, most commonly its colors, and also by our prior knowledge about its likely redshifts. To reach good photo-z performance (e.g. low bias) the prior should be strongly dependent on the selection function. Defining this in practice requires a reference sample of galaxies for which the mapping between redshift and observed properties is known, and is representative of the target galaxy selection. Hence when redshift estimates calibrated with one selection are used together with a significantly different selection during the science analysis, they are no longer guaranteed to retain their fiducial performance. (Bonnett et al., 2016; Hoyle et al., 2018).

Background galaxies ($z_{\rm src} > z_{\rm lens}$) at different redshifts (distances) contribute to the lensing signal with different amplitudes. This is characterized by the inverse of the *critical surface density*:

$$\Sigma_{\rm crit}(z_{\rm lens},\, z_{\rm src}) = \frac{c^2}{4\pi G}\frac{D_{\rm s}(z_{\rm src})}{D_{\rm l}(z_{\rm lens})\, D_{\rm ls}(z_{\rm lens}, z_{\rm src})}\,, \tag{3.1}$$

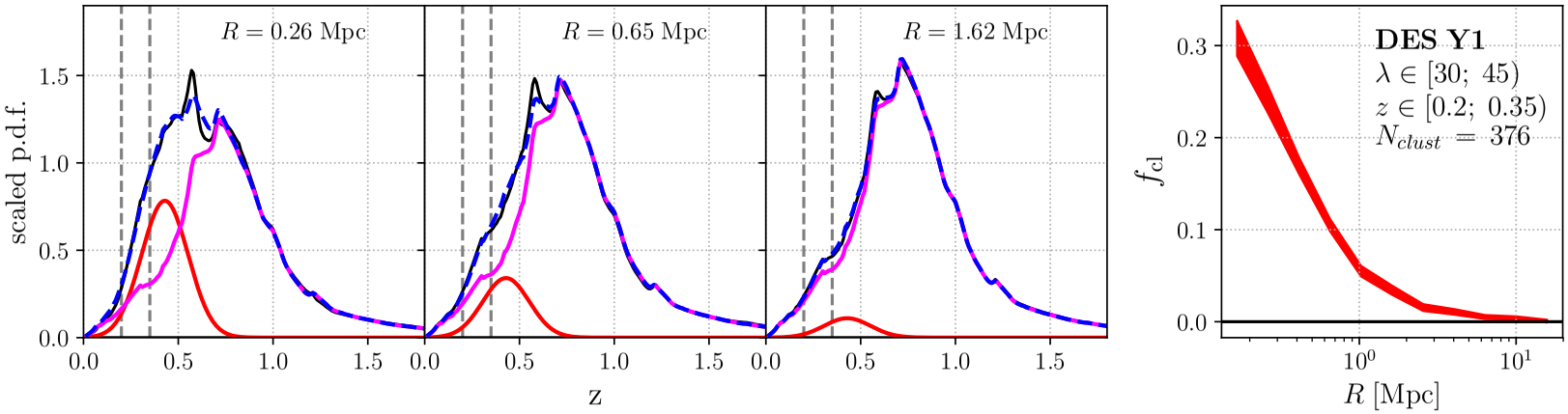

Figure 3.1: *(Left panels:)* $P(z)$ decomposition at three different radial ranges for the cluster sample with richness $\lambda \in [30;\ 45)$ and redshift $z \in [0.2; 0.35)$ in DES Y1 data. *Black lines:* average weighted $P(z\,|\,R)$ of source galaxies. *Red lines:* $P(z)$ of the Gaussian contamination component scaled by the estimated cluster member contamination rate $f_{\rm cl}$. *Magenta lines:* average $P(z\,|\,\text{field})$ scaled by $1 - f_{\rm cl}$. *Blue dashed:* model $P(z)$ calculated from the sum of the magenta and red lines. The vertical dashed lines indicate the redshift range of galaxy clusters in the cluster selection. *(Right panel:)* the cluster member contamination rate $f_{\rm cl}$ profile calculated from the decomposition presented on the left panels: the red shaded range corresponds to the amplitudes of the Gaussian components at each radial range. The width of the shaded area indicates the 1σ uncertainty region.

where D_s, D_l and D_ls denote *angular diameter* distances to the source galaxy, the lens, and between the lens and the source respectively.

In a cluster weak lensing scenario the quantity of interest is the average tangential component of the reduced gravitational shear $g_\mathrm{T} = \gamma_\mathrm{T}/(1-\kappa)$ where γ is the weak lensing shear and κ is the convergence. This is estimated from the shapes and alignments of the source galaxies through the ellipticity measure $\mathbf{e}$, where we assume that $\langle \mathbf{e} \rangle \approx \langle \mathbf{g} \rangle$. The shear signal is related to the *excess surface mass density* $\Delta\Sigma$, expressible from the physical mass distribution of the lens system via:

$$\gamma_\mathrm{T}(R) = \frac{\overline{\Sigma}(<R) - \overline{\Sigma}(R)}{\Sigma_\mathrm{crit}} = \frac{\Delta\Sigma(R)}{\Sigma_\mathrm{crit}}\,, \tag{3.2}$$

and the convergence is defined as

$$\kappa(R) = \overline{\Sigma}(R) \,/\, \Sigma_\mathrm{crit}\,, \tag{3.3}$$

where R is the projected separation from the lens, $\overline{\Sigma}(<R)$ and $\overline{\Sigma}(R)$ refer to the average surface mass density within radius R, and at radius R, respectively.

Following Sheldon et al. (2004) we define the maximum likelihood estimator for the stacked excess surface mass density of multiple clusters:

$$\widetilde{\Delta\Sigma} = \frac{\sum_i^\mathrm{src} \sum_j^\mathrm{lens} w_{i,j}\ e^\mathrm{T}_{i,j} \Big/ \langle \Sigma^{-1}_\mathrm{crit} \rangle_{i,j}}{\sum_i^\mathrm{src} \sum_j^\mathrm{lens} w_{i,j}} = \frac{\sum_i^\mathrm{src} \sum_j^\mathrm{lens} w_{i,j}\ \Delta\Sigma_{i,j}}{\sum_i^\mathrm{src} \sum_j^\mathrm{lens} w_{i,j}}\,, \tag{3.4}$$

with weights

$$w_{i,j} = \langle \Sigma^{-1}_\mathrm{crit} \rangle^2_{i,j} \Big/ \sigma^2_{e,\,i}\ , \tag{3.5}$$

Here $e^\mathrm{T}_{i,j}$ corresponds to the tangential component of the *estimated ellipticity* e of the i-th source galaxy relative to the j-th lens, $\sigma^2_{e,i}$ is the variance of the estimated shape for galaxy i, and

$$\langle \Sigma^{-1}_\mathrm{crit} \rangle_{i,j} = \int \mathrm{d}z_\mathrm{src}\ P_i(z_\mathrm{src})\ \Sigma^{-1}_{\mathrm{crit},\,i,j}(z_\mathrm{lens}, z_\mathrm{src}) \tag{3.6}$$

is defined as the *effective* inverse critical surface density for source-lens pair i,j.[1]. We note that $\Delta\Sigma$ relates to γ, however the distortion of source galaxies is determined by g, thus the effect of magnification must be accounted for during the modeling of the measured lensing signal.

The $\widetilde{\Delta\Sigma}$ estimator defined in Equation 3.4 is unbiased if the redshift p.d.f. $P(z_\mathrm{src})$ is estimated correctly. In general, intrinsic bias in the photometric redshift estimates would also bias the weak lensing measurement[2], while uncertainty alone can be propagated self-consistently (Applegate

[1] We estimate $\langle \Sigma^{-1}_\mathrm{crit} \rangle_{i,j}$, as $\langle \Sigma_\mathrm{crit} \rangle_{i,j}$ is less numerically stable.

[2] Depending on the form of the lensing estimator, it is possible for the redshift estimate to be biased in a way such that the lensing estimator is still unbiased. This however does not hold for arbitrary photo-z bias.

et al., 2014b). For this reason, weak lensing surveys spend great effort calibrating photometric redshifts for their weak lensing source galaxy catalogs (Kelly et al., 2014; Hoyle et al., 2018; Hildebrandt et al., 2017; Tanaka et al., 2018). In case the photometric redshifts are biased, it requires additional work and loss of constraining power to ensure that the photo-z bias is appropriately propagated into the systematic error budget of the final scientific results (Abbott et al., 2018).

3.2.2 Boost factor formalism

Galaxy clusters represent a large overdensity of physically associated galaxies, consisting both of actual cluster member galaxies, and also the galaxies inhabiting correlated structures. Thus cluster lines-of-sight are different from the average line-of-sight: The galaxy overdensity is concentrated at a tight peak in redshift, much narrower than the resolution of photometric redshift estimates of individual galaxies. In the transverse direction, the number density of the cluster-related galaxy population decreases as one moves away from the cluster center, where different galaxy populations (e.g. red and blue cluster galaxies) follow different radial profiles (Navarro et al., 1996; Rykoff et al., 2014).

Observationally, the weak lensing source galaxy catalog is built from galaxies selected according to morphological, photometric and spatial selection criteria (Zuntz et al., 2018; Hildebrandt et al., 2017; Mandelbaum et al., 2018), and galaxies associated with the targeted clusters also enter the catalog if they satisfy those criteria. Ideally cluster galaxies would be excluded (e.g. as in Schrabback et al., 2018a) since they are at the lens redshift and carry no lensing signal. However in a wide field survey, priors used to estimate photometric redshifts do not account for the presence of the targeted clusters. Consequently the redshift estimates can be greatly biased, and allow for cluster galaxies leaking into the source catalog with non-zero weights. Hence in the above picture the cluster member contamination can be viewed as a form of selection or representativeness bias impacting the redshift estimates. The redshift bias may further depend on galaxy type, resulting in different rates of contamination by different populations of cluster galaxies. Thus defining a high purity background sample of sufficient volume may not be possible. We note that depending on the radial separations, the contamination can originate from both the targeted galaxy clusters and also from galaxies in the correlated matter structures. For reasons of brevity we refer to both of these sources as cluster member contamination, as they can not be disentangled based purely on available redshift information.

To quantify the required boost factor correction we need to consider the impact of contamination on the $\widetilde{\Delta\Sigma}$ estimator defined in Equation 3.4. Following the method developed in Gruen et al. (2014) and extended in Melchior et al. (2017), we assume a model for the *true* line-of-sight distribution of source galaxies selected during the measurement as a combination of two terms: a cluster galaxy component which is effectively a Dirac-δ function located at $z_{\rm clust}$, and a non-cluster or background component taken to be the lensing weighted distribution of source galaxies in field lines-of-sight.

Via the above line-of-sight model, we can expand Equation 3.4 into the sum of contributions

from the cluster (cl) and background (bg) terms:

$$\begin{aligned}\widetilde{\Delta\Sigma}_{\rm est} &= \frac{\sum\limits_{j,i\in{\rm cl}} w_{i,j}\Delta\Sigma_{i,j} + \sum\limits_{j,i\in{\rm bg}} w_{i,j}\Delta\Sigma_{i,j}}{\sum w_{i,j}} \\ &= \left(\frac{\sum\limits_{j,i\in{\rm cl}} w_{i,j}}{\sum w_{i,j}}\right)\underbrace{\langle\Delta\Sigma_{i,j}\rangle_{\rm cl}}_{0} + \left(\frac{\sum\limits_{j,i\in{\rm bg}} w_{i,j}}{\sum w_{i,j}}\right)\underbrace{\langle\Delta\Sigma_{i,j}\rangle_{\rm bg}}_{\langle\Delta\Sigma_{i,j}\rangle_{\rm true}}\end{aligned} \tag{3.7}$$

of which $\langle\Delta\Sigma_{i,j}\rangle_{\rm cl}$ carries no signal, while $\langle\Delta\Sigma_{i,j}\rangle_{\rm bg}$ is the "true" signal we would estimate if there was no contamination. $\sum_{j,i\in{\rm cl}}$ and $\sum_{j,i\in{\rm bg}}$ denotes a sum over source-lens pairs with cluster members and background galaxies respectively. We furthermore define the effective contamination rate of cluster galaxies $f_{\rm cl}$ via

$$f_{\rm cl} = \frac{\sum_{j,i\in{\rm cl}} w_{i,j}}{\sum_{j,i} w_{i,j}} \ , \tag{3.8}$$

which we can use to express the boost correction needed to recover the true signal

$$\widetilde{\Delta\Sigma}_{\rm corr}(R) = \frac{\widetilde{\Delta\Sigma}(R)}{1-f_{\rm cl}(R)} \, . \tag{3.9}$$

Here $\widetilde{\Delta\Sigma}$ denotes the raw measured lensing signal obtained from Equation 3.4, $\widetilde{\Delta\Sigma}_{\rm corr}(R)$ denotes the lensing signal corrected for contamination, and $\mathcal{B} \equiv (1-f_{\rm cl})^{-1}$ is referred to as the *boost factor*. Hence in the above framework the cluster member contamination correction for a given measurement scenario is completely characterized by the $f_{\rm cl}(R)$ profile.

3.2.3 Estimating the contamination using P(z) decomposition

We estimate the contamination rate from the available color–magnitude information of source galaxies, where, due to the overdensity of the cluster we expect that the contaminating cluster galaxies will appear as a sub-population. We follow Melchior et al. (2017), and make use of the *observed* lensing-weighted average redshift probability distribution of the sources

$$P(z\,|\,R) = \frac{\sum_{i,j} w_{i,j}\, P_i(z,R)}{\sum_{i,j} w_{i,j}} \, , \tag{3.10}$$

which contain information compressed from color–magnitude space into a probability distribution function. We measure this at different projected radii R around the cluster. The weights $w_{i,j}$ for each source are identical to the ones introduced in Equation 3.4. In this framework the estimated redshifts represent information compression from the color–magnitude space into a single $P(z)$ estimate per sample. Contaminating cluster members contribute to the average photometric redshift $P(z)$-s differently in different radial ranges. Thus by tracking the changes in the $P(z)$ as a function of radius, we can recover an estimate of the underlying cluster member contamination.

We model the observed redshift distribution $P(z\,|\,R)$ as a combination of two terms, reflecting the cluster and background populations defined in section 3.2.2:

$$P(z\,|\,R) = f_{\rm cl}(R) \cdot P_{\rm memb}(z) + (1 - f_{\rm cl}(R)) \cdot P_{\rm bg}(z)\,, \tag{3.11}$$

where $P_{\rm memb}(z)$ is the redshift distribution of *contaminating* cluster member galaxies and $P_{\rm bg}(z)$ is the distribution of background galaxies (Gruen et al., 2014; Melchior et al., 2017). We approximate the second term by the appropriately weighted redshift distribution of the average survey *field* line-of-sight: $P_{\rm bg}(z) \approx P_{\rm field}(z)$. As an ansatz we consider $P_{\rm memb}(z)$ to be a Gaussian distribution. The validity of this assumption is tested in section 3.3.2.5. The free parameters of the decomposition are the mean and width of the Gaussian $P_{\rm memb}(z)$, and the contamination rate $f_{\rm cl}(R)$.

An example for this $P(z)$ decomposition method is shown on Figure 3.1 for the case of DES Y1 data. There, a qualitatively similar behavior is visible for the different radial bins, and the contamination increases with decreasing radius.

3.3 Method validation

We perform a validation benchmark to test the robustness and performance of the $P(z)$ decomposition boost estimates. First, we outline the primary assumptions of the decomposition method in section 3.3.1, validate the method in a mock analysis scenario in section 3.3.2, and perform consistency tests on DES Y1 data in section 3.3.3.

3.3.1 Model assumptions

The $P(z)$ decomposition method relies on several assumptions about both the contaminating and the background galaxies which impact the efficacy of the method. Some of these we explore below, i.e. the potential intrinsic alignment of cluster galaxies (section 3.3.1.1), the impact of weak lensing magnification (section 3.3.1.2), and the influence of blending and intra-cluster light on photometry used in the decomposition estimates (section 4.5.3). Other assumptions are tested in later sections, i.e. the Gaussian ansatz for $P_{\rm memb}(z)$ (section 3.3.2.5), and the influence of the chosen background model (section 3.3.3.3).

3.3.1.1 Intrinsic alignments

Contaminating galaxies physically connected to the lens system possess an *intrinsic alignment* due to the tidal forces acting between them. When cluster members are included in a lensing measurement, intrinsic alignments could appear as negative tangential shear around clusters due to the preferential radial orientation of galaxies. This effect is difficult to decouple from the physical lensing signal. Recent spectroscopic follow-up studies of Hao et al. (2011); Sifón et al. (2015) found no significant signal for preferential alignments of cluster member galaxies with respect to the cluster centers. Huang et al. (2018) found significant detection only when considering a high luminosity subset of galaxies, but no detection when considering their complete

galaxy sample. For this reason we assume that the dominant effect of contaminating cluster members is the dilution of the lensing signal.

3.3.1.2 Magnification

Weak lensing magnification caused by the potential of the cluster changes the observed number density and luminosity function of background galaxies. This translates into biased photometric redshift estimates for background galaxies, which would result in the $P(z)$ of background galaxies near clusters to be different from the $P(z)$ of the similarly weighted galaxies in field lines-of-sights. Gruen & Brimioulle (2017, their Appendix C), however, finds, that under realistic survey assumptions biases in photometric redshift estimates due to the increased flux of magnified sources, and due to the different surface density of magnified galaxies are sub per-cent effects that partially cancel one another.

3.3.1.3 Impact of blending and intra-cluster light

The potential bias due to blending and source obscuration in the estimated $P(z)$-s is difficult to estimate, as it would require detailed understanding of detection and shape measurement selection probabilities, as well as the photometric transfer function of representative source galaxy samples in cluster fields (see e.g. Chang et al., 2015; Suchyta et al., 2016). To a first approximation, we expect blending and source obscuration to uniformly impact all source galaxies, leading to an amplitude shift in the $P(z)$ of the selected source galaxies. However given the excellent match of the cluster background population $P(z)$ and the field background population $P(z)$ at large $z \gg z_l$ (visible in Figure 3.1), we assume that the impact of these effects is strongly subdominant. In section 3.3.3.3 we nevertheless perform a simple consistency test for differences in the background $P(z)$.

The presence of *intra-cluster light* biases the photometric redshift estimates, influencing the recovered $P(z)$-s in a manner similar to blending. However (Gruen et al., 2019, their Appendix A) estimated the impact of this effect to be negligible for the radial scales considered in this study.

3.3.2 Benchmark on the Buzzard mock observations

We test the $P(z)$ decomposition method against the true cluster member contamination in a simulated environment, mirroring the measurement setup of McClintock & Varga et al. (2019). In section 3.3.2.1 and section 3.3.2.2 we introduce the simulated observations and the mock galaxy clusters. In section 3.3.2.3 we perform the $P(z)$-decomposition on the simulated *Buzzard* data. In section 3.3.2.4 we determine the true contamination of our photo-z selected source sample. In section 3.3.2.5 we test the validity of the Gaussian ansatz for $P_{\mathrm{memb}}(z)$. Finally, in section 3.3.2.6 we discuss the agreement between the true and estimated contamination rates.

3.3.2.1 Buzzard simulated lightcones

The *Buzzard*-suite of cosmological simulations (DeRose et al., 2019) consists of mock DES Y1 catalogs generated by combining three N-body lightcones created using L-Gadget2, a version of Gadget2 (Springel, 2005) optimized for memory efficiency. The initial conditions were set up via 2nd order Lagrangian perturbation theory using 2LPTIC (Crocce et al., 2006). The lightcones were produced on the fly using simulation boxes with volumes 1050^3, 2600^3, and 4000^3 $(h^{-1}\mathrm{Mpc})^3$; the corresponding particle masses are 3.3×10^{10}, 1.6×10^{11} and 5.9×10^{11} $h^{-1}\mathrm{M}_\odot$. The resulting lightcones were joined at redshifts 0.34 and 0.9, arranged such that the highest resolution simulations are used at lower redshifts. These simulation boxes assume a flat ΛCDM cosmology with $\Omega_\mathrm{m} = 0.286$, $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_\mathrm{b} = 0.047$, $n_s = 0.96$, and $\sigma_8 = 0.82$. The galaxy catalogs were created by assigning galaxies to dark matter particles via the ADDGALS algorithm (Wechsler et al., 2018). ADDGALS calibrates the relation between the large scale density and the r-band absolute magnitudes of galaxies as measured using subhalo abundance matching (Conroy et al., 2006; Reddick et al., 2013; Lehmann et al., 2017) in a high resolution N-body simulation. For each simulated galaxy, SEDs are assigned from the SDSS DR7 VAGC (Cooper et al., 2011) by finding the galaxy in the data with the closest match in $M_r - \Sigma_5 - z$ space, where M_r is the galaxy's rest frame r-band absolute magnitude and Σ_5 is the distance to the 5th-nearest galaxy in projection (see Appendix E2 of DeRose et al., 2019). Photometric noise is added in accordance with the DES Y1 depth map of Drlica-Wagner et al. (2018), and *g,r,i,z* fluxes in the DES filters are generated from the previously assigned SEDs.

In this study we use version 1.3 of the *Buzzard* mock catalogs. Only the main "SPT"-area of DES Y1 is simulated, and the footprint is restricted to $RA < 0$ to exclude areas where the DES coverage is more inhomogeneous. This yields a simulated sky survey with a total area of 1120 deg^2. The resulting galaxy fluxes include the effect of weak lensing magnification based on ray tracing along their lines-of-sight, with the highest redshift galaxy being located at $z = 2.35$ (DeRose et al., 2019). For the purposes of the current measurement we selected sources in a way that is meant to approximate the source selection in the DES Y1 analysis (Zuntz et al., 2018) by applying S/N cuts following MacCrann et al. (2018). This sample is defined purely to mirror the properties of actual DES source galaxies, and does not contain shear or photometry systematics. We then run the BPZ template based photometric redshift algorithm (Benítez, 2000; Coe et al., 2006) on this mock catalog to obtain a $P(z)$ estimate for each source galaxy with equivalent settings as used by Hoyle et al. (2018) for the DES Y1 data. Given plausible galaxy colors and identical measurement setup, we expect BPZ to possess similar performance in Buzzard as in the DES Y1 data

3.3.2.2 Simulated galaxy clusters

In order to obtain a simulated cluster sample similar to the one presented by McClintock & Varga et al. (2019), we run the redMaPPer algorithm (Rykoff et al., 2014) on the mock galaxy catalogs with the same configuration as the real DES Y1 data.

RedMaPPer is a red-sequence based optical matched filter cluster finder which produces an estimate on the position, the optical richness λ, and redshift of the detected clusters. This yields

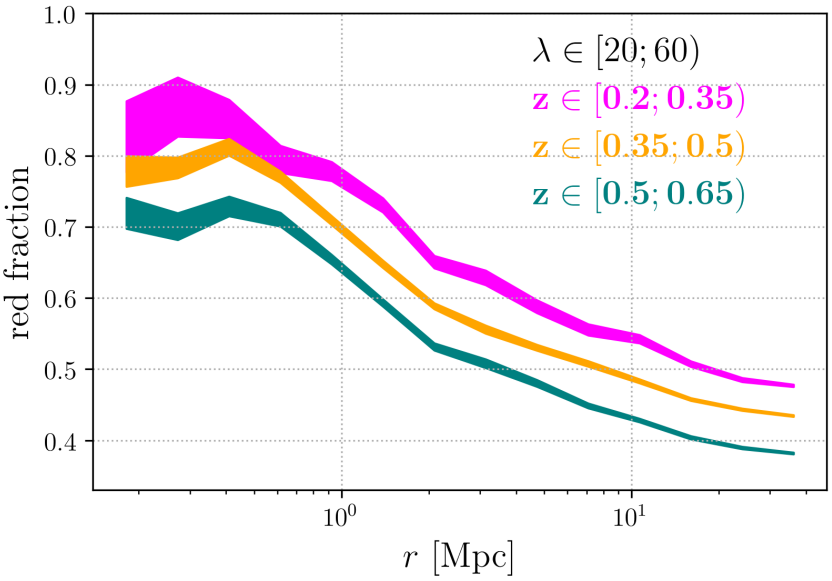

Figure 3.2: Fraction of red galaxies as a function of radius around redMaPPer clusters within the *Buzzard* mock observation across different redshift ranges but fixed richness. Shaded areas indicate 1σ statistical uncertainties estimated from Jackknife resampling.

a cluster catalog with comparable distribution in angular position and redshift to the catalog in the DES Y1 dataset. A catalog of reference random points is also generated, which are defined as positions and redshifts where the survey conditions (e.g. mask and observational depth) allow clusters of given richness to be detected.

While the redMaPPer algorithm is sensitive only to the overdensity of red-sequence galaxies, we also test the blue galaxy content of clusters as they are expected to significantly contribute to the contamination. We calculate the fraction of red galaxies as a function of radii using a rest frame magnitude limit of $m_r > -19$. For this we take a galaxy as red if it belongs to the red sequence defined in the rest frame color – magnitude space, which in practice corresponds to a cut of $(g-r) > 0.2 \cdot r - 0.028$. Figure 3.2 shows this red fraction across different redshift ranges, where we find good qualitative agreement with previous observational studies (e.g Butcher & Oemler, 1978; Hansen et al., 2009, their Figure 12). The red fractions are different across different redshift bins with a larger blue cluster member population at higher redshifts, which is expected from the time dependence of the galaxy quenching process.

A difference between the real and mock cluster catalogs is that clusters in the simulation appear to have a stronger redshift evolution in richness at a given halo mass relative to expectations from existing scaling relations. This fact along with the reduced simulated footprint results in a lower number of clusters in richness bins at low redshift compared to McClintock & Varga et al. (2019). In addition, the DES Y1 data is deeper than the reference dataset used by the ADDGALS algorithm to populate SEDs, and for this reason the mock galaxy populations and their relative abundances at faint magnitudes or high redshifts might differ from reality. Because of this, as well as because of differences in source galaxy selection and between our real and synthetic

datasets, we do not expect the cluster member contamination rates in the mock observations to be equal to our DES Y1 measurements. Nevertheless, the mocks include many qualitative aspects of the real observations, and for this reason we make use of them as a controlled environment to benchmark and validate the performance of the $P(z)$ decomposition under somewhat simplified circumstances.

3.3.2.3 Decomposition results in simulated catalogs

We estimate boost factors for the redMaPPer clusters in the mocks using an identical measurement setup as McClintock & Varga et al. (2019). Hence the $\Delta\Sigma$ estimator takes the form of:

$$\widetilde{\Delta\Sigma} = \frac{\sum \omega_{i,j} e_{\mathrm{T;i}}}{\sum \omega_{i,j} \Sigma^{\prime -1}_{\mathrm{crit;i}}} \tag{3.12}$$

with

$$\omega_{i,j} \equiv \Sigma^{-1}_{\mathrm{crit}} \left(z_{\mathrm{l}_j}, \langle z_{\mathrm{s}_i} \rangle \right) \text{ if } \langle z_{\mathrm{s}_i} \rangle > z_{\mathrm{l}_j} + 0.1 \,. \tag{3.13}$$

Where $\Sigma^{\prime -1}_{\mathrm{crit;i}}$ is calculated at a source redshift randomly drawn from the corresponding $P(z)$, while $\Sigma^{-1}_{\mathrm{crit}} \left(z_{\mathrm{l}_j}, \langle z_{\mathrm{s}_i} \rangle \right)$ represents the value at the mean redshift of the source. The mock galaxy catalog does not include shear biases thus we set the shear and selection responses to unity (**?**, **?**)see Equation 12 of][]rmy1.

Following Equation 3.7 and Equation 3.8, the contamination fraction is given by:

$$f_{\mathrm{cl}} = \frac{\sum_{\mathrm{cl}} \omega_{i,j} \Sigma^{\prime -1}_{\mathrm{crit;i,j}}}{\sum \omega_{i,j} \Sigma^{\prime -1}_{\mathrm{crit;i,j}}} \,. \tag{3.14}$$

We divide the clusters into bins of redshift $z \in [0.2; 0.4)$, $[0.4; 0.5)$, and $[0.5; 0.65)$, and richness: $\lambda \in [5; 10)$, $[10; 14)$, $[14; 20)$, $[20; 30)$, $[30; 45)$, $[45; 60)$, and $[60; \infty)$. $\widetilde{\Delta\Sigma}$ is calculated in 11 logarithmically spaced radial bins ranging from 0.2 Mpc to 30 Mpc. For each cluster sample and radial range we save $min(N_{pairs}\,;\, 2 \cdot 10^4)$ representative source-lens pairs selected in a uniform random way, and calculate the mean $P(z)$ of that source population weighted by $\omega \Sigma^{\prime -1}_{\mathrm{crit}}$. The estimate on $f_{\mathrm{cl}}(R)$ is then found by the $P(z)$ decomposition method outlined in section 3.2.3. For the field component we take the $P(z)$ of sources in the outermost radial bin, which we find to be identical to the weighted $P(z)$ of sources selected around random points in a series of Kolmogorov-Smirnov tests.

The decomposition is calculated by considering all radial scales simultaneously where we require the redshift positions and widths of the cluster components to be identical at different radial ranges. The mixing amplitudes $f_{\mathrm{cl}}(R)$ between the cluster and reference $P(z)$ are left free across radial bins. Hence the inner radial scales where the contamination is stronger provide constraints about the cluster component for the outer radial ranges. The f_{cl} profile model for a cluster sample has $N_{\mathrm{rbin}} + 2$ free parameters, and the decomposition is performed via a least squares Levenberg-Marquardt algorithm, where the optimized quantity is the mean squared deviation between the measured $P(z\,|\,R)$ and the model prediction defined in Equation 3.11. This

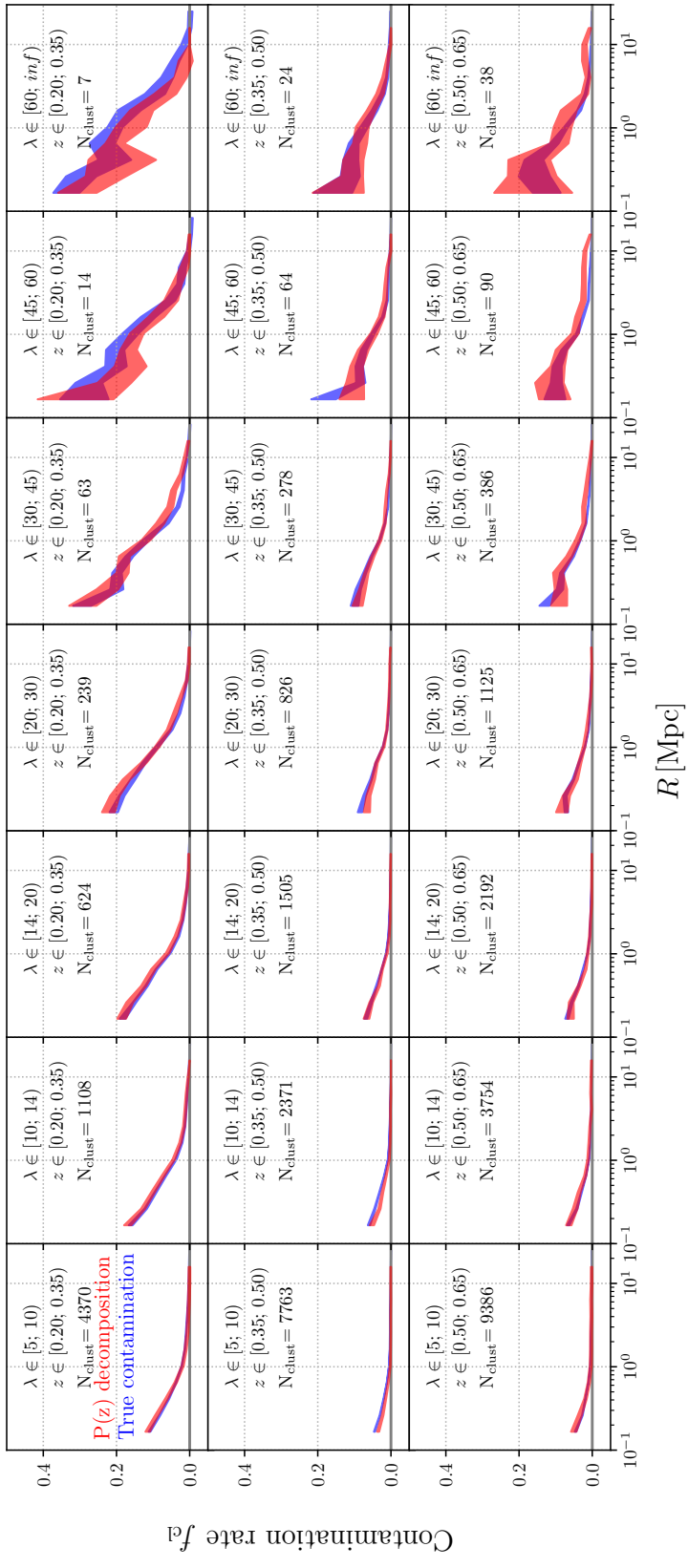

Figure 3.3: Cluster member contamination measured in the *Buzzard* mock observations for the various bins of clusters in richness λ and redshift z defined in section 3.3.2.3. *Red:* contamination profiles calculated via $P(z)$ decomposition using BPZ photometric redshift estimates with realistic photometric noise added to the mock galaxies. *Blue:* True contamination profiles calculated from the truth catalogs of the simulations via Equation 3.16.

boost factor calculation algorithm is implemented in the XPIPE python package[3], which was also used by Chang et al. (2018b) and McClintock & Varga et al. (2019), and contains an identical setup to Melchior et al. (2017).

To estimate the uncertainty on the recovered $f_{\rm cl}(R)$ we use jackknife (JK) resampling following Efron (1982):

$$\mathsf{C}^{JK}_{\widetilde{f}_{\rm cl}} = \frac{K-1}{K}\sum_{k}^{K}\left(\widetilde{f}_{{\rm cl}(k)} - \widetilde{f}_{{\rm cl}(\cdot)}\right)^{T}\cdot\left(\widetilde{f}_{{\rm cl}(k)} - \widetilde{f}_{{\rm cl}(\cdot)}\right)\,, \tag{3.15}$$

where $\widetilde{f}_{{\rm cl}(\cdot)} = \frac{1}{K}\sum_k \widetilde{f}_{{\rm cl}(k)}$ and $\widetilde{f}_{{\rm cl}(k)}$ denotes the contamination rate estimated via Equation 3.14. We make use of $K = \min\{100\,;\,N_{\rm clust}\}$ simply-connected spatial regions $\mathcal{R}_k$ for each cluster sample, defined via a spherical *k-means* algorithm[4], and $\widetilde{f}_{{\rm cl}(k)}$ is calculated from all clusters except those in region $\mathcal{R}_k$. With this method we estimate the covariance between all radial ranges within each richness and redshift bin.

The recovered contamination profiles are shown in Figure 3.3, and are qualitatively similar to those observed in the real data. The overall behavior is consistent with theoretical expectations: For all cluster bins the contamination rate decreases with increasing redshift, and a clear trend is apparent where richer clusters at a given redshift range produce greater contamination rates.

3.3.2.4 True contamination in simulated catalogs

We calculate the true contamination as the *excess* rate for galaxies to be located within the immediate $\mathcal{S}_j \equiv [z_j - \Delta z;\ z_j + \Delta z]$ vicinity of the clusters, defined via

$$f_{\rm cl}^{\rm true} \equiv \frac{\sum_{j}^{N_c}\sum_{z_{s,i}\in\mathcal{S}_j}\omega_{i,j}\Sigma'^{-1}_{\rm crit;i,j}}{\sum_{j}^{N_c}\sum_{i}^{N_s}\omega_{i,j}\Sigma'^{-1}_{\rm crit;i,j}} - \frac{\sum_{l}^{N_r}\sum_{z_{s,i}\in\mathcal{S}_l}\omega_{i,j}\Sigma'^{-1}_{\rm crit;i,l}}{\sum_{l}^{N_r}\sum_{i}^{N_s}\omega_{i,j}\Sigma'^{-1}_{\rm crit;i,l}}\,, \tag{3.16}$$

where N_c refers to the number of clusters, N_r to the number of random points, and N_s to the number of source galaxies, while $\omega_{i,j}$ is the lensing weight associated with the source-lens pairs defined in Equation 3.13. $\mathcal{S}_j$ and $\mathcal{S}_l$ refer to the immediate true redshift vicinities of clusters and random points respectively. That is, $f_{\rm cl}^{\rm true}$ is the probability of finding a galaxy within the redshift range $\mathcal{S}$ around the clusters, minus the same probability for random lines-of-sight, where the second term we obtain by saving source-lens pairs around redMaPPer random points. This is equivalent to a cylindrical selection of contaminating galaxies, which is motivated by the fact that the contamination originates not only from physically bound galaxies, but also from galaxies in the extended correlated structures.

Figure 3.4 illustrates the above approach. It is clear that a large fraction of source-lens pairs near cluster centers (*blue*) actually lie at the cluster redshift. Comparing this with the distribution

[3] https://github.com/vargatn/xpipe

[4] https://github.com/esheldon/kmeans_radec

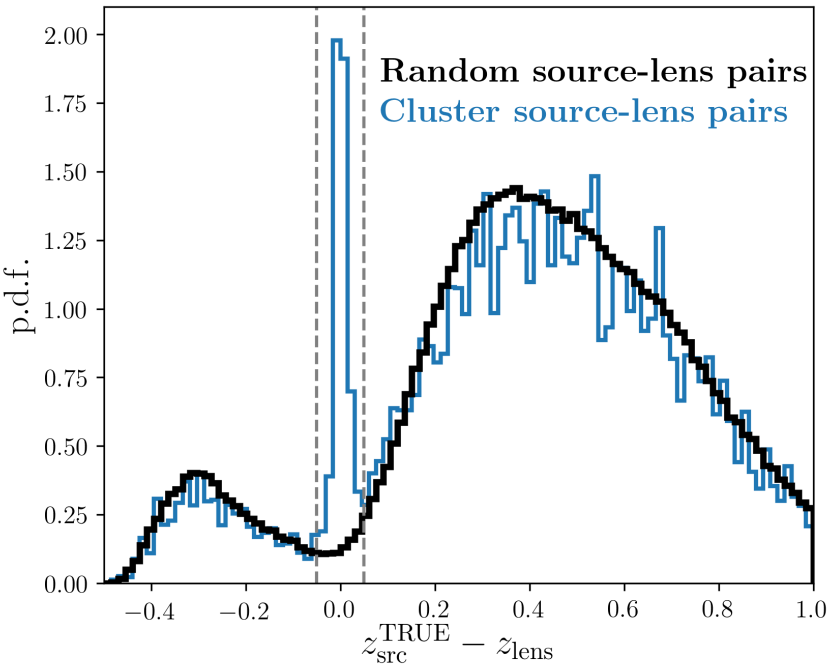

Figure 3.4: Schematic for estimating the *true* cluster member contamination fraction in the *Buzzard* mock observations. The figure shows histograms of the weighted, true redshift separation of source lens-pairs at different radial distances from galaxy clusters with $z \in [0.35; 0.5)$ and $\lambda \in [30; 45)$. *Blue:* source-lens pairs at low radial scales around clusters ($R < 0.78\,Mpc$). *Black:* source–lens pairs around redMaPPer selected random points in the same radial range. *Gray dashed:* $\Delta z = \pm 0.05$ vicinity of the cluster redshift.

of galaxies around random points (*black*), the contamination rate is taken as the excess area under the curve within the $\pm\Delta z$ (*dashed*) region. Based on Figure 3.4, we adopt $\Delta z = 0.05$ as our fiducial redshift width for the purposes of computing the true contamination rate. The resulting $f_{\rm cl}^{\rm true}$ profiles are shown on Figure 3.3 as the blue shaded regions, where the 1σ uncertainties are estimated from jackknife resampling using the same approach as in section 3.3.2.3.

3.3.2.5 Validity of Gaussian cluster model

With the formalism introduced in section 3.3.2.4, we can compute the average, weighted, photometric redshift $P(z)$ for likely cluster galaxies which are located within the $\pm\Delta z^{\rm TRUE}$ vicinities of clusters. As shown on Figure 3.5 these $P(z)$-s have a prominent peak located slightly above the cluster redshift range. Cluster galaxy $P(z)$-s also possess a long tail extending up to high redshifts. This is an intrinsic feature of photometric redshift estimation, as some cluster galaxies have spectral types that exhibit these types of degeneracies.

Figure 3.5 also shows the best fit Gaussian cluster component models $P_{\rm memb}(z)$. These are obtained from the decomposition method in section 3.3.2.3 and are not informed of the true cluster member $P(z)$-s. Due to the chosen analytic form of $P_{\rm memb}(z)$, the long high redshift tail

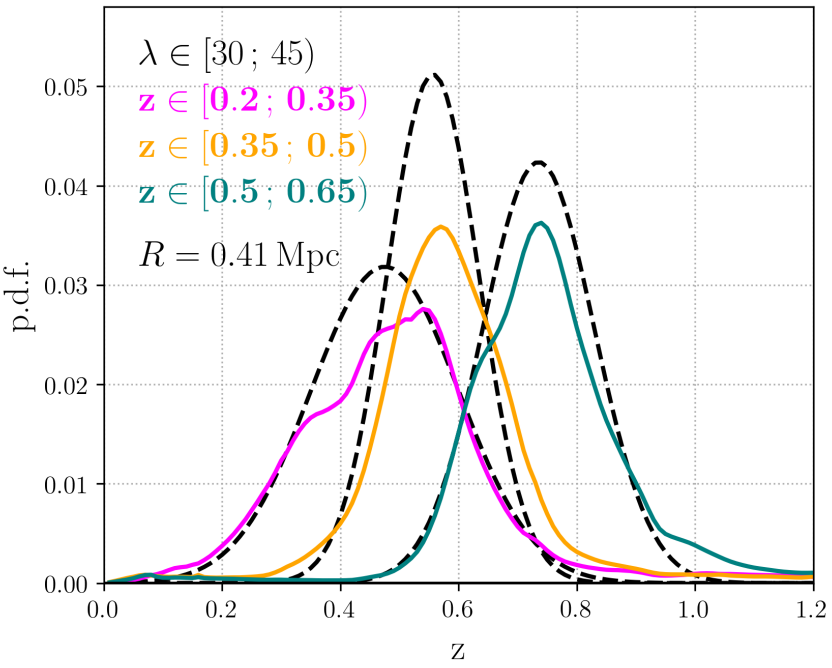

Figure 3.5: Comparison of the Gaussian $P(z)$ component model with the actual $P(z)$-s of likely cluster galaxies in the *Buzzard* mock observations. *Colored curves:* $P(z)$ of source galaxies within $\pm\Delta z^{\mathrm{TRUE}}$ of the clusters, in different cluster redshift bins, but at the same richness and radial bin. *Dashed curves:* Best fit curves of the Gaussian cluster component model $P_{\mathrm{memb}}(z)$. Note that the shown P(z)-s are normalized for $z \in [0;\, 3.5]$.

of the actual cluster galaxy $P(z)$-s can not be recovered, which results in the apparent offset of the (normalized) p.d.f-s on Figure 3.5. For the shown samples the best fit Gaussian contains within 1σ 57%, 47% and 56% of probability of the actual $P(z)$ of cluster members. Nevertheless, the Gaussians recover the approximate position and width of the peaks, and possesses fewer degrees of freedom than alternative high-skewness models.

3.3.2.6 Discussion of simulation benchmarks

Figure 3.3 compares the boost factor profiles estimated from $P(z)$ decomposition from simulations as described in section 3.3.2.3, with the actual cluster member contamination rate calculated in section 3.3.2.4. We find that our estimated boost factors are in excellent agreement with the true member contamination rates in the *Buzzard* mock simulations. The uncertainties shown in Figure 3.3 are estimated via Jackknife resampling, and do not incorporate *systematic* uncertainties. Hence we estimate this systematic uncertainty by requiring consistency between the true and estimated f_{cl} profiles across all parameter bins with $N_{\mathrm{clust}} > 50$. Via this approach we find a global relative systematic uncertainty of $\delta_{sys} < 1$ percent across different richness – redshift selections, where the total covariance is given by $\mathsf{C}_{i,j} = \mathsf{C}^{JK}_{i,j} + \delta_{i,j} \cdot f^2_{\mathrm{cl};\,i} \cdot \delta^2_{sys}$

We note that the simulated galaxy catalogs include the effects of magnification with the typi-

cal angular resolution of 0.6 arcminutes, corresponding to approximately 0.15, 0.2 and 0.24 Mpc in the different redshift bins. While this low resolution allows for only weak constraints, the good agreement between the estimated and true $f_{\rm cl}$ profiles indicate that magnification does not play a significant role in the resolved radial ranges.

The purpose of this simulation benchmark is to test how well the $P(z)$ decomposition predictions match the contamination within the simulation, not to extrapolate for the real DES data. Thus we do not require full realism from the simulated environment. Nevertheless, the simulated color distribution of galaxies has been studied by DeRose et al. (2019) in a setup nearly identical to version 1.3 of Buzzard used in the present study. They found the simulated galaxy properties to be broadly consistent with reality except for a slight systematic shift on the color of the blue cloud. We do not anticipate that this manifests in a qualitative difference on the performance of the boost factor estimator compared to real data. We however note that the presented mock scenario is constructed to resemble the DES analysis setup, including the choice of photometric redshift estimation algorithm. For this reason the observed performance is not necessarily indicative for a significantly different survey and analysis scenario. While the abundances, radial profiles, and color properties of cluster galaxies in the simulation may be slightly different from reality, we expect Buzzard to be qualitatively similar to the real DES Y1 data. Hence we take the excellent performance of the $P(z)$ decomposition in this setting as a strong motivation for its applicability for real observations.

3.3.3 Analysis on DES Y1 Data

In this section we apply the $P(z)$ decomposition method to DES Y1 data, following the exact measurement setup presented in McClintock & Varga et al. (2019). The structure of this section is the following: in section 3.3.3.1 we present the relevant parts of the DES Y1 dataset relating to the galaxy cluster catalog and the weak lensing source galaxies, in section 3.3.3.2 we derive the form of the necessary boost factor correction, while in section 3.3.3.3 present a simple test on the robustness of our contamination model, and finally in section 3.3.3.4 we compare with the alternative method of correlation based boost factor estimate.

3.3.3.1 The DES Y1 dataset

The DES Y1 observations cover approximately 1800 deg^2 of the southern sky in *g,r,i,z* bands. These observations are processed via a variety of photometric data reduction steps into the Y1 GOLD catalog (Drlica-Wagner et al., 2018) which is the main science quality catalog of DES. Using the fiducial multi-epoch, multi-object fitting algorithm (MOF) DES finds the 10σ limiting magnitudes of this dataset for 2" apertures to be $g \approx 23.7$, $r \approx 23.5$, $i \approx 22.9$ and $z \approx 22.2$. Based on these observations McClintock & Varga et al. (2019) defined a locally volume limited catalog of galaxy clusters identified via the redMaPPer algorithm. In the Y1 footprint the average MOF limiting magnitude is deep enough to detect a 0.2 L^* galaxy up to $z \approx 0.7$, thereby setting the maximum depth of the volume limited cluster sample.

Approximately 1500 deg^2 of this catalog is further processed by the METACALIBRATION algorithm (Huff & Mandelbaum, 2017; Sheldon & Huff, 2017) to define a source galaxy sample

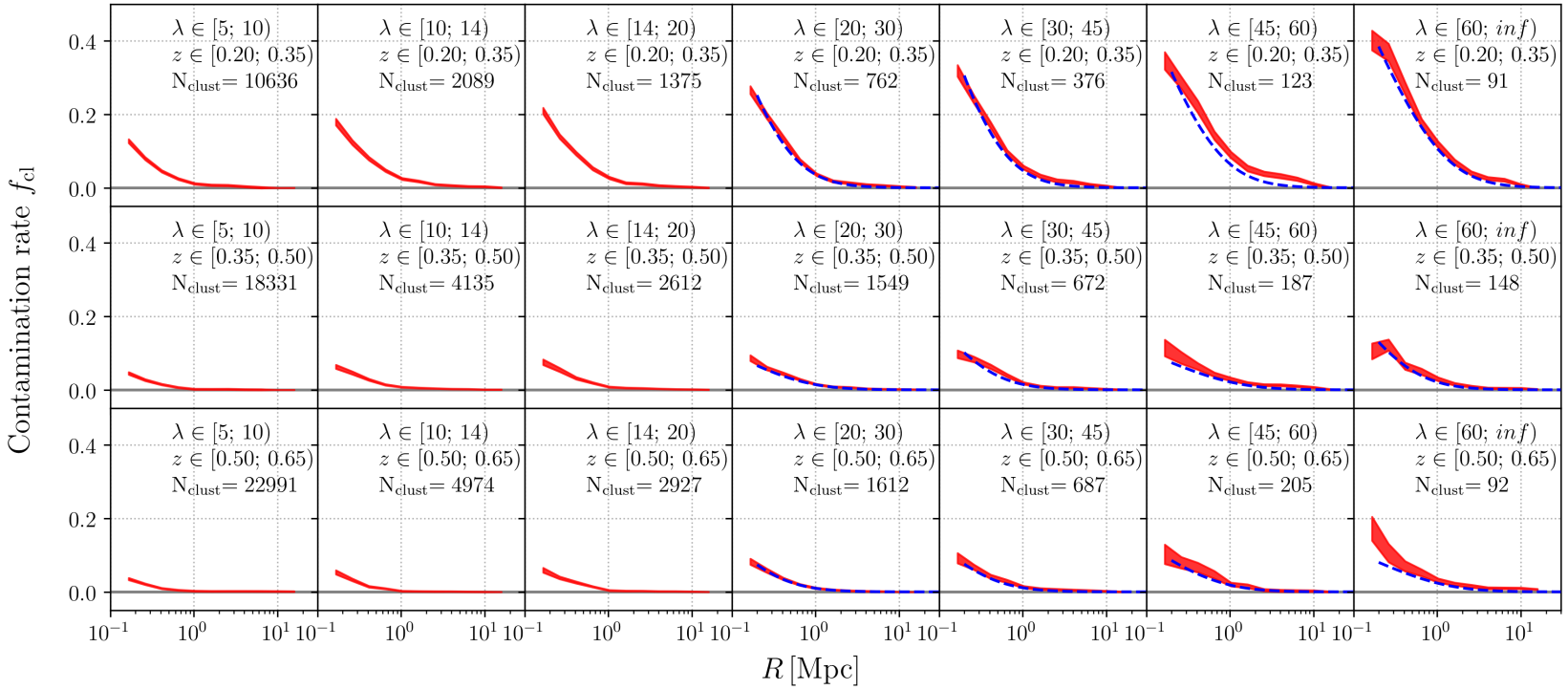

Figure 3.6: Cluster member contamination measured in the DES Y1 data for the various bins of clusters in richness λ and redshift z defined in section 3.3.3.2. *Red:* contamination profiles calculated via $P(z)$ decomposition using BPZ redshift estimates. *Blue:* Boost factor model from section 3.4.2. The curves correspond to the best fit parameters found by McClintock & Varga et al. (2019) from a likelihood optimization performed jointly with the cluster mass reconstruction.

(Zuntz et al., 2018). This source galaxy catalog consists of an ellipticity estimate e_i for each galaxy, along with ancillary quantities used to perform the bias calibration via the *response* $\mathsf{R} = \mathsf{R}_\gamma + \mathsf{R}_{\rm sel}$ of the ellipticity estimates to shear and the source galaxy selection function respectively.

Photometric redshift $P(z)$-s are calculated via the BPZ template based algorithm (Hoyle et al., 2018). Two separate redshift estimates are derived: one based on the MOF-based galaxy colors listed in the GOLD catalog, and a second based on the photometric model obtained from METACALIBRATION.

Hoyle et al. (2018) found the derived redshift estimates to be mildly biased in the mean redshift, which McClintock & Varga et al. (2019) corrected by introducing a multiplicative calibration to the the final weak lensing measurements. Weak lensing measurements are dominantly sensitive to the mean redshift of the source dn/dz, while the $P(z)$ decomposition method is sensitive not to absolute shifts, bit differences in relative shape of $P(z)$-s. Due to this, as the expected effect of a shift is estimated to be negligible we do not explicitly apply such redshift bias correction.

3.3.3.2 Contamination estimator for DES Y1-like data

In this section we derive the required boost factor correction for the $\Delta\Sigma$ estimator employed for the DES Y1 analysis. For this, Equation 3.12 and Equation 3.13 are replaced with:

$$\widetilde{\Delta\Sigma} = \frac{\sum \hat{\omega}_{i,j} e_{\mathrm{T;i}}}{\sum \hat{\omega}_{i,j} \mathsf{R}^{\mathrm{T}}_{\gamma;\,i,j} \Sigma'^{-1}_{\mathrm{crit;\,i,j}}} \,. \tag{3.17}$$

where

$$\hat{\omega}_{i,j} \equiv \Sigma^{-1}_{\rm crit} \left(z_{\mathrm{l}_j}, \langle z^{\rm MCAL}_{\mathrm{s}_i} \rangle \right) \text{ if } \langle z^{\rm MCAL}_{\mathrm{s}_i} \rangle > z_{\mathrm{l}_j} + 0.1 \,, \tag{3.18}$$

which is the general form of the estimator. Here we neglected the selection response term, which McClintock & Varga et al. (2019) found to be subdominant compared to the shear response R^T_γ, where the superscript refers to the response matrix rotated into the tangential frame. In the above estimator the weighting and selection is performed based on the mean METACALIBRATION based redshift estimates $\langle z^{\rm MCAL}_{\mathrm{s}_i} \rangle$, while $\Sigma'^{-1}_{\mathrm{crit;\,i,j}}$ is calculated using a random draw from the MOF-based redshift $P(z)$.

Following Equation 3.7 and Equation 3.8 we find the contamination rate to be:

$$f_{\rm cl} = \frac{\sum_{cl} \hat{\omega}_{i,j} \mathsf{R}^{\mathrm{T}}_{\gamma;\,\mathrm{i,j}} \Sigma'^{-1}_{\mathrm{crit;\,i,j}}}{\sum \hat{\omega}_{i,j} \mathsf{R}^{\mathrm{T}}_{\gamma;\,\mathrm{i,j}} \Sigma'^{-1}_{\mathrm{crit;\,i,j}}} \,. \tag{3.19}$$

We perform the $P(z)$ decomposition in a setup identical to section 3.3.2.3, but using weights according to Equation 3.19, and make use of a randomly selected, representative subsample of the source-lens pairs from McClintock & Varga et al. (2019). The detailed description of our results is presented in section 3.4, while the boost profiles themselves are shown on Figure 3.6.

3.3.3.3 Sensitivity to background component choice

The performance of the $P(z)$ decomposition method is dependent on how well the ansatz for the background component resemble the p.d.f. of actual background galaxies. Furthermore, the average $P(z)$ estimated for a galaxy sample may contain minor features (e.g. wiggles and peaks) which depend on the internal setup of the photometric redshift algorithm (e.g. distribution of templates within the BPZ algorithm), and do not themselves relate to the physical distribution of galaxies (Bonnett et al., 2016; Melchior et al., 2017). Differences originating from these non-physical reasons may also impact the robustness of the contamination estimates.

We test the self-consistency of the decomposition method and its sensitivity to the minor features in the estimated $P(z)$-s by extending the fiducial $P(z)$ decomposition analysis with a second step. In this second step the reference $P(z)$ model is updated from the "field" $P(z)$ to the observed $P(z)$ at $R \approx 1$ Mpc minus the Gaussian cluster model found in the previous step. The $f_{\rm cl}$ fit is then repeated with this new reference $P(z)$ component, while keeping the position and width of the Gaussian cluster model component unchanged. The resulting boost factor profiles are shown on Figure 3.7, overlaid with the fiducial boost factor profiles. The two iterations agree very well, and following the approach used in section 3.3.2.6 we estimate a relative systematic uncertainty of < 1 percent, motivating that the choice for the background component propagates to only a negligible difference in the final contamination profiles.

3.3.3.4 Comparison with correlation based boost factors

An alternative way for estimating boost factors is via the angular clustering of source galaxies around clusters, as only the contaminating galaxies are correlated with the cluster (Sheldon et al., 2004; Applegate et al., 2014b; Hoekstra et al., 2015; Simet et al., 2017b; Leauthaud et al., 2017). We calculate this correlation function via the estimator:

$$Corr = \frac{N_R}{N_D} \cdot \frac{DD}{RR} - 1 \,, \tag{3.20}$$

where DD and RR are defined as $\sum \omega \mathsf{R}_\gamma^{\rm T} \Sigma_{\rm crit}^{\prime -1}$ around redMaPPer clusters and random points respectively, while N_D refers to the number of clusters, and N_R to the number of random points (Landy & Szalay, 1993). The results of this measurement are shown on Figure 3.7. The correlation function estimates are, for many cluster samples, preferentially lower than the $P(z)$ decomposition estimates, especially at the two lower redshift selections.

This can be understood as clusters impacting the spatial distribution of source galaxies in ways other than contamination by cluster galaxies: e.g. the density and blending of cluster members may lead to a bias against selecting sources near clusters (Simet et al., 2017a; Leauthaud et al., 2017; Zuntz et al., 2018), which can explain the preferential lower estimates. Such effects cannot be captured by random points, as they relate to the presence of the cluster in the line-of-sight, and are not well characterized for the DES Y1 METACALIBRATION shear catalogs. By contrast, the $P(z)$ decomposition method is insensitive to color-agnostic fluctuations in the source selection, and to the number density profile of source galaxies.

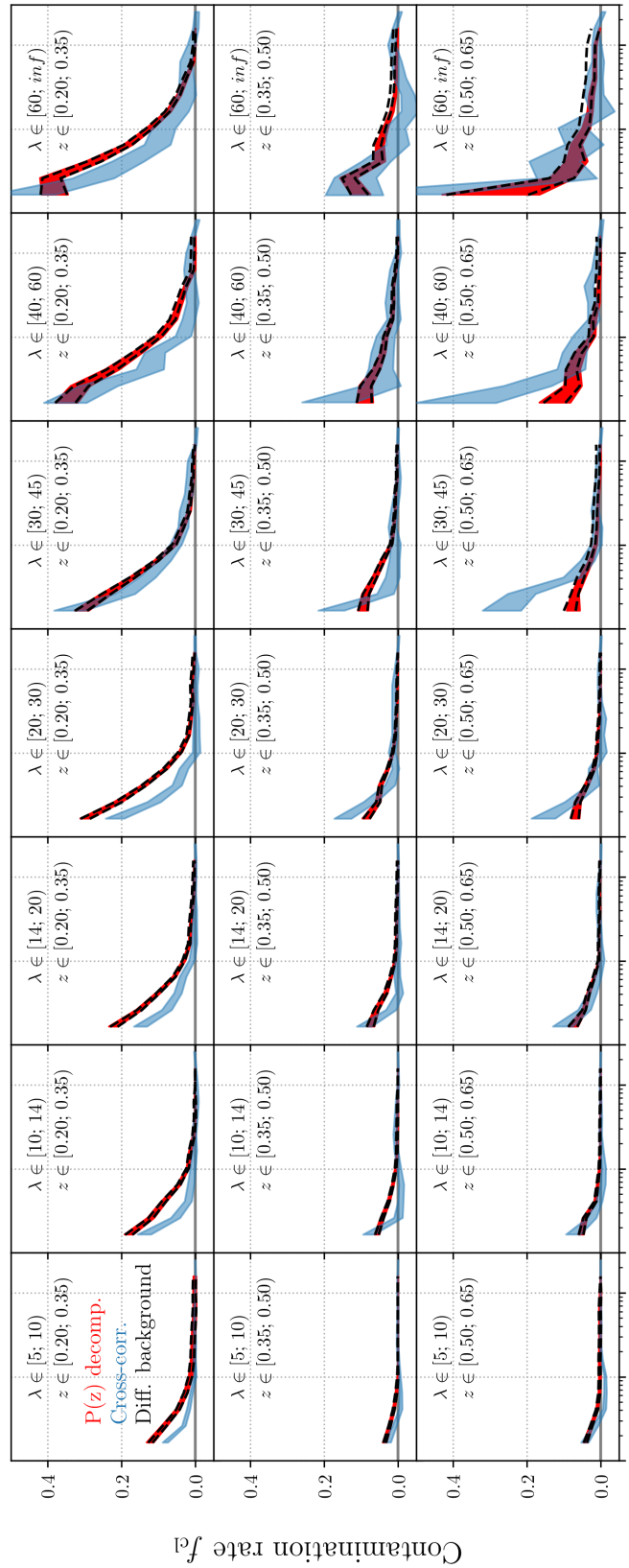

Figure 3.7: Comparison of the $P(z)$ decomposition based contamination estimates *(red)* with the cross-correlation based estimates *(blue)* across the different richness and redshift bins. *Black:* results from the background sensitivity test. The *black* background sensitivity results are found to agree very well with the fiducial red curves, while the *blue* correlation based estimates appear to be globally biased low.

3.4 Results for DES Y1 data

3.4.1 Boost factor estimates

We present the contamination rate estimates from applying our method to the DES Y1 data in Figure 3.6. In the present calculation we consider all cluster selections, but note that in McClintock & Varga et al. (2019) only the $\lambda > 20$ clusters enter the determination of the mass–observable relation. The qualitative behavior of the contamination rate profiles agrees well with the theoretical expectation of decreasing contamination with increasing radius. As expected, the amplitude of the contamination increases with cluster richness. Furthermore, the contamination rates are higher for low-redshift clusters, as for those fainter cluster member galaxies can also be detected, whose photometric redshifts are less accurate.

We find that the "peak" in the $P(z)$ due to contaminating galaxies is very prominent at low radii for all cluster bins, and the presence of this feature is critical for the applicability of the decomposition method. The best fit parameters of the Gaussian $P_{\rm memb}(z)$ model are presented in Figure 3.8, along with the used prior ranges. The means of these best-fit Gaussian $P_{\rm memb}(z)$ distributions differ from the redshift ranges of the clusters. However this is expected from the way source galaxies are selected in the DES analysis: only those cluster member galaxies enter the source selection whose estimated mean redshift scatters towards higher redshifts.

The contamination rate profiles $f_{cl}(R)$ shown in Figure 3.6 can be directly translated into a multiplicative correction factor $\mathcal{B} \equiv (1 - f_{\rm cl})^{-1}$ necessary for recovering an unbiased estimate on $\Delta\Sigma$ via Equation 3.9.

3.4.2 Analytic boost factor model

We model the boost factor profile using a Navarro-Frenk-White (NFW) profile (Navarro et al., 1996):

$$\mathcal{B}_{\rm model}(R) = 1 + B_0 \frac{1 - F(x)}{x^2 - 1}\,, \tag{3.21}$$

where $x = R/R_s$, and

$$F(x) = \begin{cases} \frac{\tan^{-1}\sqrt{x^2-1}}{\sqrt{x^2-1}} & : x > 1 \\ 1 & : x = 1 \\ \frac{\tanh^{-1}\sqrt{1-x^2}}{\sqrt{1-x^2}} & : x < 1 \end{cases}\,. \tag{3.22}$$

This model has two free parameters per cluster bin: B_0 and R_s which characterize the amplitude and scale radius of the correction profile respectively.

The best fit boost model profiles are overlayed on Figure 3.6 to the raw contamination rate estimates. Following the approach used in section 3.3.2.6 we estimate a relative systematic uncertainty of < 1 percent. In McClintock & Varga et al. (2019) these fits are performed in a joint likelihood analysis together with the mass profile model and systematic corrections. This way the estimated statistical uncertainty of the boost factors is propagated self-consistently into their final mass constraints. The model parameters are not tied to the mass parameters of clusters

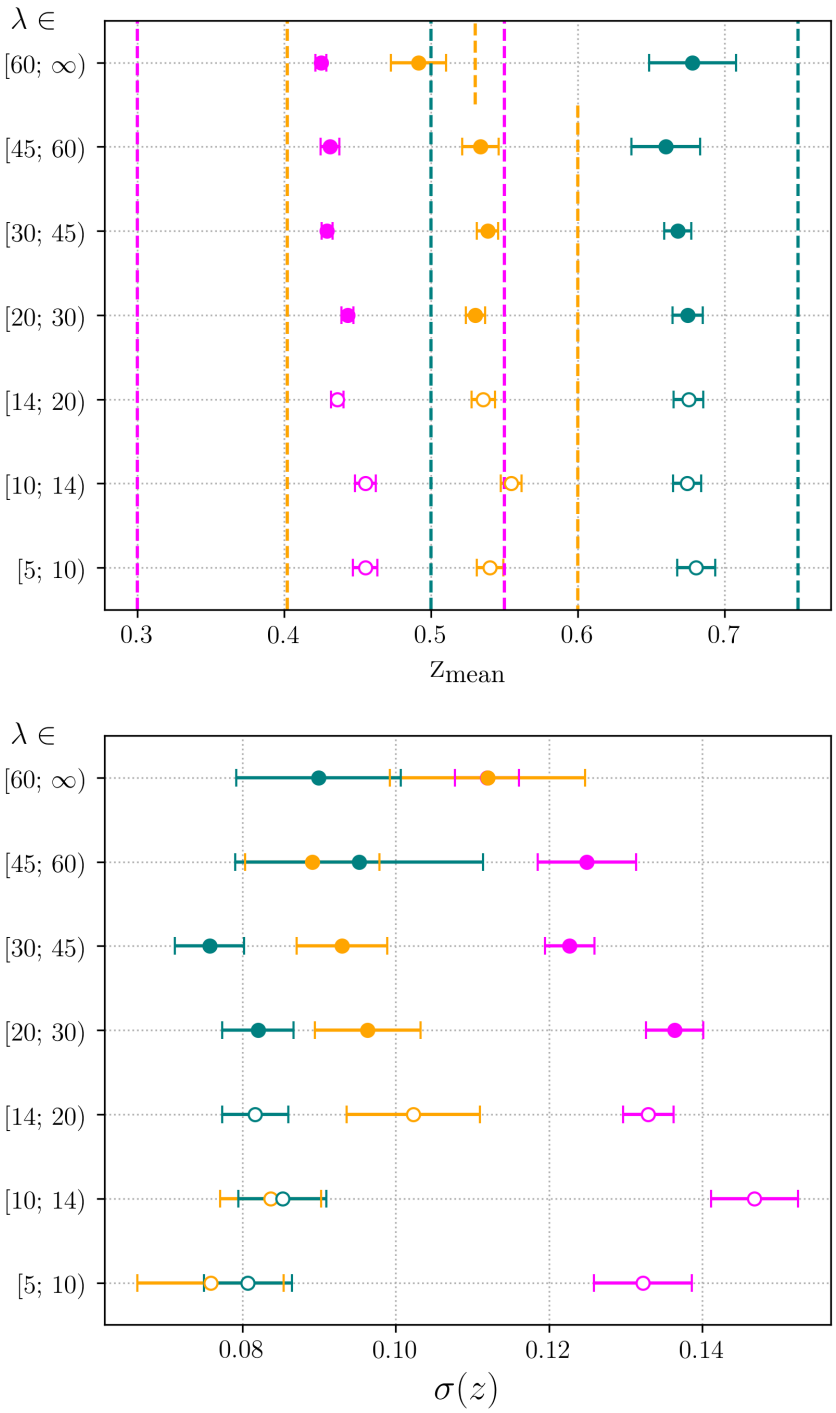

Figure 3.8: Best fit parameters for the Gaussian $P_{\rm memb}$ component found for the $P(z)$ decomposition in the DES Y1 data. The colors indicate the various cluster redshift bins: $z \in [0.2, 0.35)$, $z \in [0.35, 0.5)$ and $z \in [0.5, 0.65)$ is denoted by *magenta*, *orange* and *green* respectively. *(Top panel:)* mean redshift of the cluster member components. The dashed colored lines indicate the prior range for the mean of the cluster component. *(Bottom panel:)* standard deviation of the cluster member components.

to allow for freedom in describing the boost factors. Due to the excellent performance of the $P(z)$ decomposition method in our tests, and as the systematic uncertainty found in a simulated environment was subdominant compared to the Jackknife error estimate, McClintock & Varga et al. (2019) did not assume any additional systematic uncertainty to this source of systematic error.

We note that in the Monte Carlo chains run by McClintock & Varga et al. (2019) the R_s and B_0 parameters were found to be degenerate, hence the increase in actual contamination does not translate into an obvious increase in B_0. However this was found to propagate into only a mild change in the boost factor profile over the studied radial range. The recovered cluster masses were robust against degeneracy in the boost factor model parameters, and are not significantly impacted. Nevertheless, we find that the scale radius of the contamination component is typically at least twice as large as the scale radius of the NFW halo. This is consistent with the expectation that the galaxy distribution of clusters can be described as an NFW distribution with lower (approximately half) concentration than the underlying dark matter halo (Budzynski et al., 2012).

3.5 Summary and Conclusions

In this study we carried out a detailed method validation on the $P(z)$ decomposition cluster member contamination estimation algorithm proposed by Gruen et al. (2014) and Melchior et al. (2017). This approach relies on the decomposition of the average redshift $P(z)$-s of source galaxies around galaxy clusters into a cluster member and background component, to obtain an estimate on the relative number of contaminating galaxies which are mistakenly included in the source galaxy catalog. Since its inception this method has been used by studies ranging from the DES Science Verification cluster mass calibration (Melchior et al., 2017), to cluster weak lensing studies focusing on the detection of the *splashback*-feature (Chang et al., 2018b), and to the mass calibration of SPT selected clusters (Stern et al., 2018). It also serves as an important constituent of the weak lensing mass calibration on DES Y1 (McClintock & Varga et al., 2019), which will be used in deriving cosmological constraints based on the number counts of optically identified galaxy clusters (DES Collaboration et al., 2020).

In order to demonstrate the applicability of $P(z)$ decomposition based boost factors we performed a series of tests benchmarking various aspects of the approach. We find the following:

- The method performed well in a mock survey simulation (section 3.3.2), yielding excellent agreement between the estimated contamination rates and the actual true number of contaminants extracted from the truth catalogs of the simulation.

- Within the mock analysis we investigated the validity of the Gaussian ansatz for the cluster $P(z)$ component (section 3.3.2.4). We found that it recovers the approximate redshift and width of the peak within the $P(z)$ of the contaminating galaxies. Furthermore the Gaussian ansatz did not appreciably bias the estimated contamination.

- We tested the sensitivity of the contamination estimates to the choice of the background

$P(z)$ component on DES Y1 data (section 3.3.3.3), and found an excellent agreement between the boost factors derived via the fiducial and alternative background components.

- We compared the method with an alternative, transverse correlation based contamination estimate in section 3.3.3.4. We found indications that the alternative method is preferentially underestimating the contaminations, which is likely an imprint of the radial source galaxy selection function. This is expected to impact the decomposition predictions to a lesser extent as it does not make use of the number profile of sources.

Excluding galaxies from the source catalog can also reduce the cluster member contamination, however it may also reduce the statistical power of the measurement if the exclusion criteria are too broad. Hence it presents a trade-off in the total error budget between the systematic uncertainty originating from boost factors and statistical uncertainty such as shape noise. However based on the consistency tests presented in this paper, and on the fact that boost factors played a strongly subdominant role in the total error budget of McClintock & Varga et al. (2019) it appears that the $P(z)$ decomposition method is sufficient to provide boost factor estimates for current cluster weak lensing analyses. We note that this determination is dependent on the characteristics of the sky survey e.g. depth, area, number of filters. Notably cluster weak lensing studies such as Medezinski et al. (2018d,b); Miyatake et al. (2018) in the ongoing Hyper Suprime-Cam Survey (Aihara et al., 2018) favored the approach of trying to excluding cluster member galaxies via color - color or $P(z)$ cuts.

During the DES Y1 analysis we propagated uncertainties by making use of a simple analytic model – an NFW profile – to describe the boost factor correction. Previously Melchior et al. (2017) also used an analytic model, while others such as Chang et al. (2018b) and Stern et al. (2018) chose to directly use the recovered boost factor profiles in correcting their $\Delta\Sigma$ measurements. While the NFW model was found to be a sufficient description for the analysis of McClintock & Varga et al. (2019), it is likely that with the increasing precision of future studies more complex boost factor models might become necessary.

We quantified several possible sources of systematic uncertainty impacting the $P(z)$ decomposition method, finding < 1 percent relative systematic uncertainty based on benchmarks on mock observations, < 1 percent relative systematic uncertainty originating from the choice of the background $P(z)$ component, and < 1 percent relative systematic uncertainty from requiring good global agreement between the numerical boost factor estimates and the analytic model. From these contributions we estimate that the decomposition method under optimal circumstances can provide boost factor estimates with approximately 2 percent relative global systematic uncertainty. However we note that specific circumstances such as the performance of the photometric redshift algorithm, or the source galaxy selection function will impact the accuracy and precision of the $P(z)$ decomposition method.

Chapter 4

Synthetic Galaxy Cluster Observations Based on DES Y3 Data

Scientific context The previous chapters of original research (chapter 2 and chapter 3) focused on developing a cluster mass measurement method for the DES Year 1 dataset. The aim of those analyses was to estimate the mean cluster mass via weak lensing for a population of clusters within a bin of observables, which is broadly composed of two major tasks:

1. *Measure the lensing profile.* This includes preparing the input dataset of source galaxies with shape and photometric redshift estimates and performing the stacked weak lensing measurement to obtain $\Delta\Sigma(R)$ or $\gamma_{\mathrm{T}}(\theta)$ and a corresponding covariance estimate. Furthermore, this step includes the characterization and correction for photometric redshift and shape estimation biases influencing the the lensing profile.

2. *Estimate cluster masses based on the lensing profile.* This includes defining the mass model, applying the necessary corrections, performing the likelihood analysis, and finally estimating the mean cluster mass $\langle M \rangle$ for the parameter bin. Furthermore additional calibrations relating to cluster mis-centering, selection effect and model bias corrections are also performed at this stage.

Of these, step 2 is closely related to the selection function used in defining the cluster sample, while step 1 is largely related to the dataset and sky survey.

The task we now embark on aims to investigate the performance of weak lensing profile measurements (step 1). During this next chapter of original research we aim to construct a synthetic test environment for galaxy cluster weak lensing, by creating mock observational data products mimicking those in real cluster lensing measurements. The produced mock observations will allow future studies to perform a complete end-to-end reproduction of a weak lensing analysis including the photometric analysis from survey-like images, shape and photometric redshift estimation, and the measurement of the lensing profile using a setup identical to the real DES analysis, and to propagate the various systematic biases and noise terms in precision weak lensing measurements into a joint uncertainty estimate.

Authorship and publication The research presented in this chapter is conducted within the Dark Energy Survey, and is currently being finalized for submission to internal review within the collaboration, before the eventual submission to the peer-reviewed scientific journal (Varga et al., in prep).

The single leading author of this study is T. N. Varga (myself), however we thank the advisory contribution of D. Gruen, and S. Seitz. Furthermore we thank the help of N. McCrann, E. Sheldon and M. Jarvis with adapting the DES image simulation pipeline for use within this study. The kernel density estimation and bandwidth selection problem for galaxy populations in DES has been also explored in the Master's thesis of C. v. Wedemeyer which was co-supervised by myself, S. Seitz and R. Bender, and we acknowledge the useful discussion and insight gained from that work.

4.1 Introduction

The study of galaxy clusters has in recent years become a prominent pathway towards understanding the nonlinear growth of cosmic structure, and towards constraining the cosmological parameters of the universe (Kravtsov & Borgani, 2012; Weinberg et al., 2013). Weak gravitational lensing provides a practical method to study the mass properties of clusters, as it is directly connected to the gravitational potential of the targets, and it is readily scalable to a statistical ensemble of targets in wide field surveys. For this reason the lensing based mass calibration of galaxy clusters became a standard practice for galaxy cluster based cosmological analyses (Rozo et al., 2010; Mantz et al., 2015; Planck Collaboration et al., 2016; Bocquet et al., 2019).

Weak lensing measurements rely on estimating the shapes or ellipticities of background *source* galaxies, which enables inferring a quantity called gravitational shear γ. The shear field is directly related to the gravitational potential of the targeted (lens) object, and by estimating the preferential orientation of source galaxies one can learn about the projected mass distribution within the lens system (Bartelmann & Schneider, 2001). In practice, there are multiple approaches in estimating the ellipticity of a galaxy. Commonly used methods include model fitting and measurements of second moments, with several innovative approaches developed in recent literature (Zuntz et al., 2013; Refregier & Amara, 2014; Miller et al., 2013; Bernstein & Armstrong, 2014b; Huff & Mandelbaum, 2017; Sheldon & Huff, 2017).

Irrespective of the chosen family of algorithms, the performance of the shear estimates cannot be a-priory guaranteed, and needs to be validated in a series of tests (Jarvis et al., 2016; Zuntz et al., 2018; Samuroff et al., 2018). This validation commonly involves the creation of mock observations, *image simulations* which are then processed and used to estimate the bias and uncertainty of the different methods in a controlled measurement scenario (Massey et al., 2007; Bridle et al., 2009; Mandelbaum et al., 2015; Samuroff et al., 2018; Kannawadi et al., 2019; Pujol et al., 2019). Furthermore the bias of lensing measurements may vary within the survey footprint due to to blending with other light sources both fully resolved and unresolved, differences in galaxy properties (e.g. morphology, shear level) and the local selection function of the catalog. This is especially important in the case of galaxy clusters, which present a significant deviation from the cosmic median line-of-sight due the overdensity of cluster member galaxies (Hansen et al., 2009; To et al., 2019), their diffuse intra-cluster light (ICL) (Zhang et al., 2019a; Gruen et al., 2019; Kluge et al., 2020) and the characteristically stronger shears induced in their vicinity (see section 2.4.1.1). These considerations pose a requirement to extend the validation of the weak lensing estimators to cluster weak lensing scenario.

The aim of this study is to create synthetic observations of galaxy clusters, which enable performing a full-scale mock cluster weak lensing analysis in a controlled survey-like environment. Our approach is to construct a statistical learning algorithm to generate plausible mock galaxy clusters and model their lines-of-sight in a way which captures the real distribution of galaxies and their observed photometric feature (e.g. luminosity, color, morphology) distributions with respect to the properties of the targeted clusters. These mock clusters and lines-of-sight are then rendered into images in the same format as actual survey observations and can further processed with the standard data reduction and analysis pipelines of the survey.

The essence of our statistical learning algorithm is to estimate the radius dependent color

distribution of cluster member galaxies by comparing the radial color distribution around galaxy clusters and a set of reference random points in a relatively shallow wide field survey. We further use this distribution as a prior for the galaxy features of fainter galaxies and their joint distribution measured from a well characterized small area deep field survey. This approach transforms the deep field distribution to resemble the feature distribution of cluster member galaxies and field galaxies respectively. It keeps the additional features measured accurately in the deep field data, and extrapolating the luminosity function of the mock line-of-sight to fainter magnitudes. Galaxy clusters are therefore reconstructed jointly with the line-of-sight galaxies using more precise and accurate photometric solutions from the deep field data, and inheriting features not directly measured in the wide field, e.g. additional bands, morphology parameters or redshifts. This approach shortcuts the computational cost and limited representation of reality of numerical simulations. The creation of a mock line-of-sight with galaxy positions and properties is formulated as a random draw from a feature distribution which enables efficient scalability for generating large numbers of mock cluster realizations required for benchmarking precision measurements.

This work is done in preparation of a cluster weak lensing benchmark analysis for the Dark Energy Survey, and therefore aims in its scope to focus on its particular observational details. Nevertheless the presented algorithm is expected to be easily generalized to other upcoming surveys such as LSST (LSST Science Collaboration et al., 2009) and Euclid (Amendola et al., 2018) as well. The structure of this chapter is the following: First in section 4.2 we introduce the DES year 3 dataset, in section 4.3 we outline the statistical approach used in modeling the synthetic lines-of-sight, in section 4.4 describe the concrete results of the galaxy distribution models derived from the DES Y3 dataset, and finally in section 4.5 outline the method for generating mock observations for DES Y3.

In the following, unless otherwise indicated, we assume a flat ΛCDM cosmology with $\Omega_\mathrm{m} = 0.3$ and $H_0 = 70$ km s^{-1} Mpc^{-1}, with distances defined in physical coordinates, rather than comoving.

4.2 DES Y3 Data

At present, the Dark Energy Survey (DES) is processing and analyzing the first three years of observations made between 2013 August 15 and 2016 February 12 (DES Collaboration et al., 2016; DES Collaboration, in prep a). This Year 3 (Y3) dataset has achieved nearly full footprint coverage albeit at shallower depth, with on average 4 tilings in each band out of the eventually planned 10 tilings. From the full 5000 deg^2, the effective survey area is reduced to approximately 4400 deg^2 due to the masking of the Large Magellanic Cloud and bright stars. In parallel to the wide field survey a smaller, deeper supernova deep field survey is also conducted covering 27 deg^2 in 4 patches. These consist of in total 10 un-dithered pointings of the Dark Energy Camera (DECam, Flaugher et al., 2015) repeated on a weekly cadence, reaching resulting in a deep field survey 1.5 - 2 mag deeper than the wide field survey.

The DES Y3 footprint is shown in Figure 4.1. Due to data availability, of the four deep field patches only three is considered in the present study. These consist of 8 partially overlapping tilings and are denoted as SN-C, SN-E and SN-X, and their location is also shown on Figure 4.1.

In the following we introduce the different DES Y3 data products used in the chapter, and refer the reader for section 1.4 for a more fundamental introduction to the DES survey strategy. We note that these photometric catalogs and the galaxy cluster catalog will serve as the basis of the fiducial weak lensing cosmology analyses performed within the scope of the DES Y3 effort.

4.2.1 Wide Field Data

The science quality photometric catalog of DES Y3 is the Y3A2 GOLD dataset (DES Collaboration, in prep a). This includes catalogs of photometric detections and parameters from the wide field survey as well as the corresponding maps of the characteristics of the observations, foreground masks, and star-galaxy classification.

Data processing starts with single epoch images for which de-trending and photometric corrections are applied. They are subsequently co-added to facilitate the detection of fainter objects. The base set of photometric detections is obtained via SExtractor (Bertin & Arnouts, 1996) from $r + i + z$ coadds. The fiducial photometric properties for these detections are derived using the *single-object-fitting* (SOF) algorithm based on the `ngmix` (Sheldon, 2015) software which performs a simultaneous fit of a bulge + disk model to multi-epoch exposures while for each exposure, modeling the point spread function (PSF) as a Gaussian mixture. An expansion of this model is the *multi-object-fitting* (MOF) (Drlica-Wagner et al., 2018; Zuntz et al., 2018) approach where in addition to the above first step friends-of-friends (FoF) groups of galaxies are identified based on their fiducial models, and in a subsequent step the galaxy models are corrected for all members of a FoF group in a combined fit. While for the Y3A2 GOLD dataset the SOF and MOF photometry were found to yield similar solutions, it is expected that in crowded environments due to its more advanced treatment of blending the MOF photometry would perform better.

The fiducial depth of the Y3A2 photometric catalog is characterized as the 10σ limiting magnitude for $2''$ apertures which are found to be $g = 24.28$, $r = 23.95$, $i = 23.34$ and $z = 22.63$ (DES Collaboration, in prep a). We note that all magnitudes used in the present analysis are defined in the AB system. When considering the 10σ detection limit for galaxies in SOF photometry the depth is reduced to $g = 23.78$, $r = 23.56$, $i = 23.04$, $z = 22.39$. As an alternative measure of survey depth the detection completeness of galaxies is estimated on the overlap between DES Y3 wide and deep field surveys, where it is found that the SOF galaxy catalog is 99% complete for galaxies with $i < 22.5$. Star - galaxy separation is performed based on the morphology derived from SOF and MOF quantities, which for the $i < 22.5$ sample has 98.5% efficiency and 99% purity, yielding approximately 226 million extended objects out of a base sample of 390 million detections.

SOF and MOF derived magnitudes are corrected for atmospheric and instrumental effects using a forward modeling approach, a spectral energy distribution (SED) based chromatic correction on a per object level, and a correction for interstellar extinction to obtain the final corrected magnitudes MAG_CORRECTED.

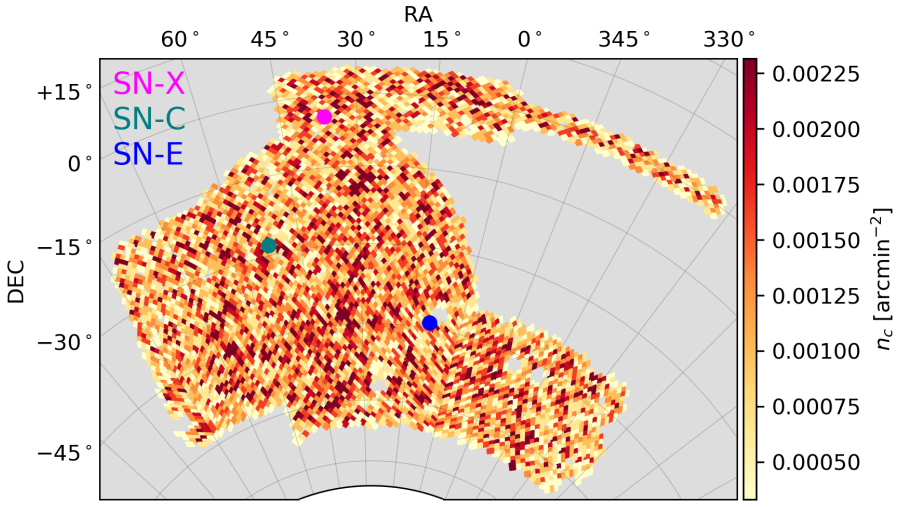

Figure 4.1: Footprint of targeted clusters in DES Y3. Blue markers: location of Deep field regions COSMOS, SN-C, SN-E, SN-X (marker size not to scale)

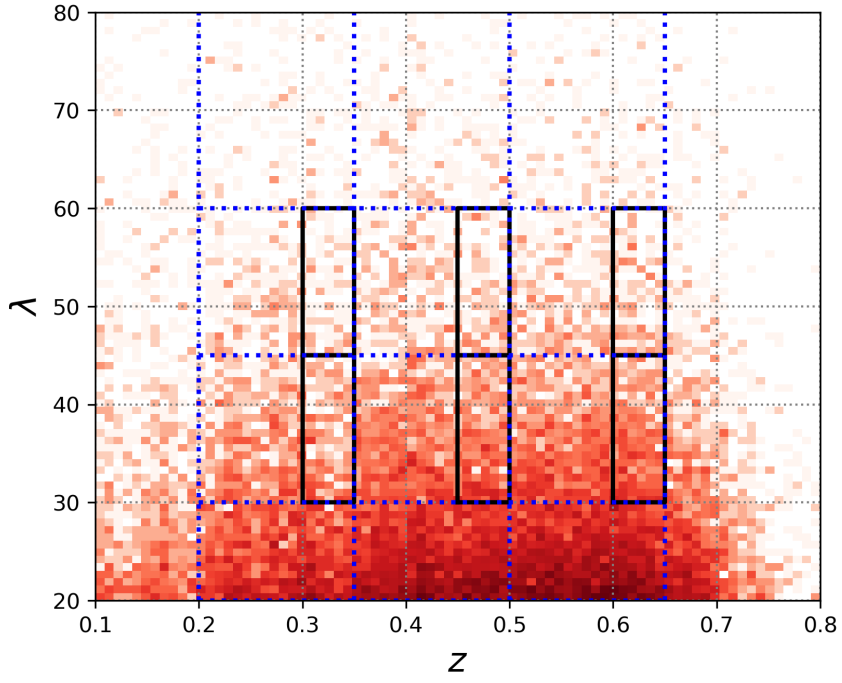

Figure 4.2: Distribution of redMaPPer clusters in DES Y3 dataset in the volume limited sample. *solid black:* narrow redshift selection, *blue dotted*: DES Y1 cluster cosmology selection

4.2.2 RedMaPPer Cluster Catalog

We consider an optically selected sample of galaxy clusters identified by the redMaPPer algorithm in the DES Y3 data. The base input for this cluster finding is the Y3A2 SOF photometry catalog described above, from which redMaPPer identifies galaxy clusters as overdensities of red-sequence galaxies. The redMaPPer version used in DES Y3 is v6.4.22+2, which is equivalent to v6.4.17 (used in chapter 2 for the DES Y1 catalog) with minor updates. For a more detailed introduction we refer the reader to section 1.4.1 and section 2.2.2.

Two cluster catalogs are defined: a flux-limited and a locally volume limited sample in a similar manner to section 2.2.2. For the present study we consider the latter, as that sample assumes no extrapolation of the cluster luminosity function during the cluster identification. For DES Y3 the survey limiting magnitude is deep enough to image a $0.2\,L_*$ galaxy at $z \approx 0.65$, which corresponds to $i \approx 22.55$, thereby setting the redshift limit for the volume limited cluster catalog.

The volume limited redMaPPer cluster catalog contains more than 869,000 clusters down to $\lambda > 5$ and more than 21,000 above $\lambda > 20$. The spatial distribution of the latter higher richness sample is shown on Figure 4.1, and the richness and redshift distribution is shown on Figure 4.2. In addition to the cluster catalog, a catalog of reference random points is also provided which are drawn from the part of the footprint where survey conditions permit the detection of a cluster of given richness and redshift.

4.2.3 Deep Field Data

The supernova (SN) and deep field survey is a concurrent side survey of DES which was followed up on a roughly weekly cadence in u, g, r, i, z, Y bands taking up approximately 30% of the total DES observation time (DES Collaboration, in prep b). The deep field survey is organized into four distinct supernova fields: SN-S, SN-X, SN-C and SN-E. These areas are divided into a total of 10 DECam pointings resulting in a deep field footprint of 27 deg^2, which in contrast with the wide field survey are not dithered to improve survey depth and PSF modeling.

The deep field observational campaign serves a two-fold scientific motivation: enabling the DES SN cosmology program (Abbott et al., 2019), and providing testing ground for benchmarking the photometric performance of the DES wide field survey. Furthermore the overlap of the deep fields with external multi-band data from the VIDEO survey (Jarvis et al., 2013) for fields: SN-X, SN-C, SN-E enables the improvement and benchmarking of photometric redshift estimates for wide field objects, by providing J, H, K band coverage. These overlap fields consists of 8 of the 10 DECam pointings. In addition to the above DES also covered the COSMOS field (Scoville, 2003), however due to the different observation strategy that is not part of the present analysis.

The deep field data processing strategy differs from the wide field in that here in addition to the DES observing program, images from community DECam observations are also included. This results in significantly more single exposure images than planned for in the fiducial survey. Together with the depth of the DES data, this allows for multiple co-addition strategies aimed for different science cases. In the present study we consider only the detections derived from

the COADD_TRUTH stacking strategy which aims to to optimize for reaching approximately $10\times$ the wide field survey depth while requiring that the deep field resolution (FWHM) be no worse that the median FWHM in the wide field data (DES Collaboration, in prep b). The deep field detection and photometric processing is done jointly between DES observations and the corresponding external survey data from VIDEO which are re-analyzed to arrive to a joint photometric solution. A key difference compared to Y3A2 GOLD is that the MOF algorithm is run with "forced photometry" where astrometry and deblending is first done using DECam data, and only in a second step are infrared bands incorporated while keeping the detections and morphology of the DECam solutions. For the resulting MOF magnitudes photometric corrections are applied in the same manner as for the Y3A2 GOLD dataset.

Due to availability, in the present analysis we only consider the MOF catalog from the SN-X, SN-C and SN-E fields. For these fields (DES Collaboration, in prep a) performed detailed consistency tests comparing the deep field derived stellar properties with expected galactic stellar populations and comparing the photometry solutions between the deep field and the re-analyzed wide field observations over the same footprint, finding very good agreement.

As an experimental data product, a photometric redshift estimate is provided by DES Collaboration (in prep a) for the deep field galaxies via the `EAzY` algorithm (Brammer et al., 2008). These photometric redshift estimates are obtained by fitting a mixture of stellar population templates to the *ugrizJHK* band fluxes of the deep field galaxies. The possible galaxy redshifts and stellar template parameters are varied jointly to obtain a redshift probability distribution. In the case of photometric redshift estimation, estimating the redshift p.d.f is only the first step, and the resulting redshifts need to be validated and corrected for biases (Hoyle et al., 2018). This validation step usually takes significant additional effort, and was not performed by the DES collaboration for the `EAzY` catalog at the time of the writing of this study. Nevertheless, as the algorithm presented in this analysis does **not** rely on photometric redshifts, but considers them only as an additional feature of galaxies which was measured in the deep field survey, we include the `EAzY` photometric redshift estimates in the deep field input catalog. They however serve only as a placeholder quantity, and we caution the reader to only consider them as such.

4.3 Statistical Model

The essence of this study is to measure and model the average galaxy content of redMaPPer selected galaxy clusters within a cluster sample and to use this measurement to create mock galaxy clusters. These mock clusters are constructed to represent the typical cluster line-of-sight, and do not aim to capture cluster-to-cluster or line-of-sight to line-of-sight variations.

By construction, the clusters identified by redMaPPer are always centered on a bright central galaxy (BCG). Central galaxies form a unique and small subset of all galaxies, and therefore we treat them separately from non-central galaxies. In our synthetic observations we always consider a mock central galaxy with the mean redMaPPer BCG properties to correspond to the mean cluster. Thus the task for the rest of this section is left as to model the properties and distribution of non-central galaxies (in the following simply denoted as galaxies).

Our approach is formulated in two steps: First, measuring and modeling galaxy properties

and distributions. Secondly based on the model, generate mock lines-of-sight consisting of simulated galaxies, and render them into survey-like observations mimicking the exposures taken by real telescopes, while applying the appropriate weak lensing effects. The resulting mock observations then allow us to benchmark cluster weak lensing measurements in a controlled test scenario.

Throughout this analysis we assume that galaxies are to first order sufficiently described by a set of observable features, primarily provided by the DES photometric processing pipeline. The key features are: i-band magnitude m_i with de-reddening and other relevant photometric corrections applied, multi-band colors $\vec{c} = (g-r,\ r-i,\ i-z)$, galaxy redshift z_g, and morphology parameters $\vec{s}$ describing the scale radius, ellipticity and flux ratio of the two components of the `ngmix` SOF/MOF bulge + disk galaxy model. The full list of features and their relation to the DES Y3 data products is listed in Table 4.1

Our aim is to model the distribution of galaxies in the space of the above features as a function of projected separation R from galaxy clusters of richness λ and redshift z. This distribution cannot be directly measured from the DES wide field survey, as the bulk of the galaxy population lies beyond the completeness threshold magnitude of $i \approx 22.5$, where photometric errors come to dominate the derived features. To counteract this limitation we adopt a two-step approach: First a target distribution of reference colors and radius $(\vec{c}_{ref};\ R \,|\, \lambda, z)$ is measured in the wide field survey. In the second step the wide field target distribution is used as a prior for resampling the galaxy features measured in the DES deep fields. Comparing the target distribution around clusters and around a set of reference random points enables us to isolate the feature distribution of cluster members. Thus the resampling can transform the deep field feature distribution into an estimate on the full feature distribution of cluster member galaxies it can transfer the additional features measured accurately only in the deep field data, and extrapolate the cluster population to fainter magnitudes.

Figure 4.3 shows an illustration of a mock cluster generated as a result of this analysis at the level of a galaxy catalog and also as a fully rendered DES Y3-like coadd image, along with an actual redMaPPer cluster taken from the DES Y3 footprint with similar richness and redshift.

The flowchart of the algorithm used in the statistical model of the clusters is shown in Figure 4.4, which introduces the various components shown is the purpose of this section. The structure is therefore the following: in section 4.3.1 we describe the measurement of the radial feature correlation function around galaxy clusters and random points (`Radial Correlator`), in section 4.3.2 we introduce the methodology of the various feature distributions as kernel density estimates (`KDE Pipeline`), in section 4.3.3 we outline the statistical background subtraction and survey extrapolation method implemented in `Survey Resampling`, and finally in section 4.3.4 describe the full algorithm for isolating the cluster member feature distribution.

4.3.1 Photometric Feature Correlation Function

We group galaxy clusters into two bins of richness $\lambda \in [30; 45)$ and $[45; 60)$, and three bins of redshift $z \in [0.3; 0.35)$, $[0.45; 0.5)$ and $0.6; 0.65)$ where each sample is processed separately. Our binning scheme is motivated by the selections of chapter 2 which served as the basis for DES Y1 cluster cosmology DES Collaboration et al. (2020). In this pathfinder study however we only

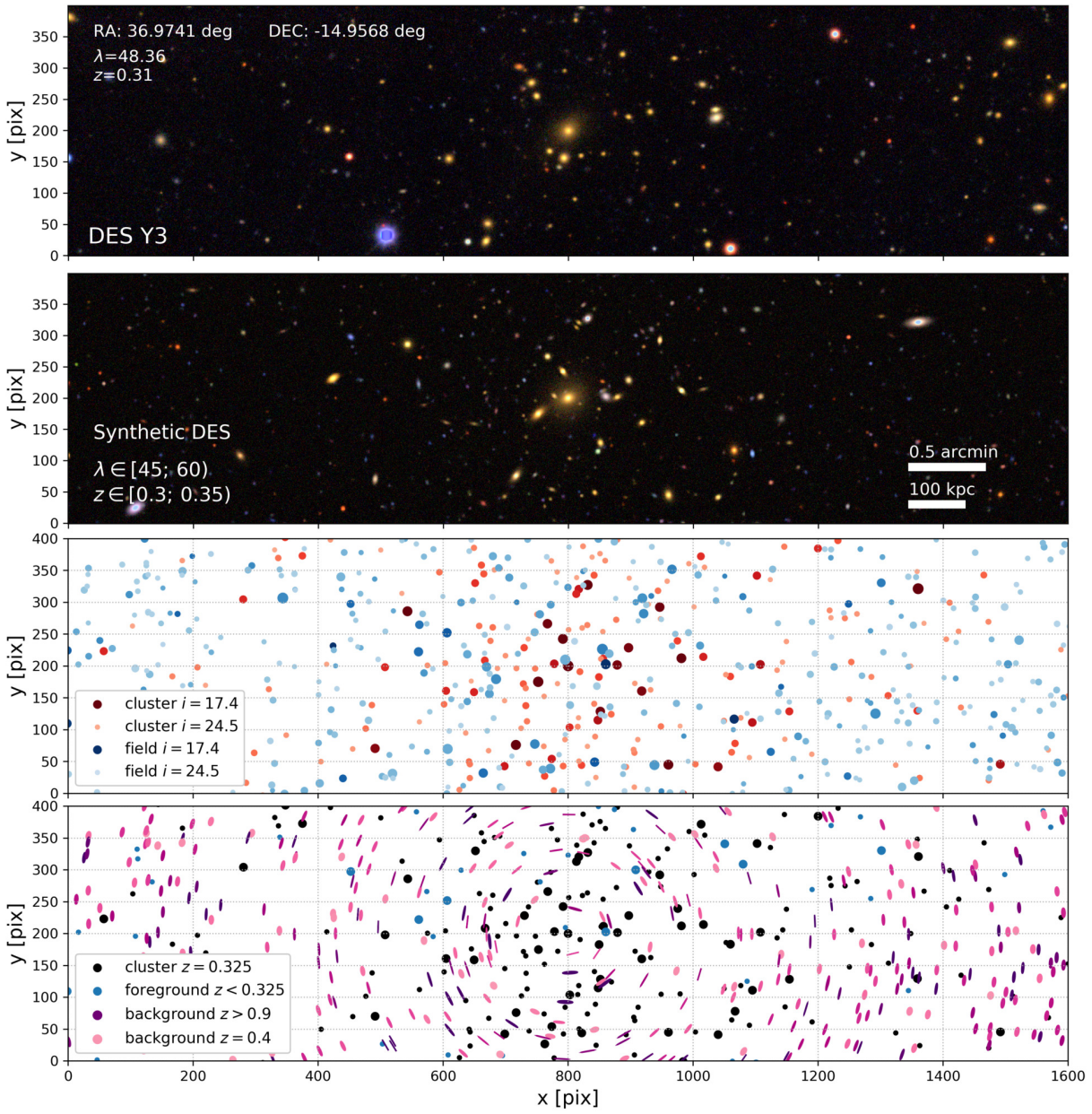

Figure 4.3: Real and synthetic galaxy cluster side by side. *Top: gri* color composit image of a real redMaPPer galaxy cluster in the DES Y3 footprint. *Second row: gri* color composit image of a synthetic galaxy galaxy cluster representative of $\lambda \in [45\ 60)$, $z \in [0.3; 0.35)$. *Third row:* brightness distribution of the synthetic light sources for cluster members (red/brown) and foreground and background objects (blue). Darker shades and larger symbols correspond to brighter objects. *Bottom row:* exaggerated shear map of background sources (magenta ellipses) with the shade representing redshift, cluster members (black) and foreground sources (blue).

Feature	catalog parameter	description
Deep field features		
m	bdf_mag_dered_3	i-band MOF magnitude with photometric correction
$\vec{c}$	bdf_mag_dered_2 - bdf_mag_dered_1	g - r MOF color with photometric correction
	bdf_mag_dered_3 - bdf_mag_dered_2	r - i MOF color with photometric correction
	bdf_mag_dered_4 - bdf_mag_dered_3	i - z MOF color with photometric correction
$\vec{s}$	sqrt(bdf_g_0^2 + bdf_g_1^2)	absolute MOF ellpiticity
	FRACDEV	bulge / disk flux ratio at fixed component size
	$\log_{10}$(1 + bdf_T)	MOF size squared in arcsec^2 $T =< x^2 > + < y^2 >$
z_g	z_mc	*ugrizYJHK*-band based photo-z estimate from EAzY
Wide field features		
R	$\log_{10}\sqrt{(\mathrm{RA} - \mathrm{ra_{ref}})^2 + (\mathrm{DEC} - \mathrm{dec_{ref}})^2}$	$\log_{10}$ projected separation in arcmin from reference point
m	MOF_CM_MAG_CORRECTED_I	i-band MOF magnitude with photometric correction
$\vec{c}$	MOF_CM_MAG_CORRECTED_G - MOF_CM_MAG_CORRECTED_R	g - r MOF color with photometric correction
	MOF_CM_MAG_CORRECTED_R - MOF_CM_MAG_CORRECTED_I	r - i MOF color with photometric correction
	MOF_CM_MAG_CORRECTED_I - MOF_CM_MAG_CORRECTED_Z	i - z MOF color with photometric correction

Table 4.1: Features and their definitions from the column of the relevant photometric catalogs. Deep field features: DES Y3 deep and supernova fields (DES Collaboration, in prep b) for further explanation see section 4.2.3. Wide field features: DES Y3 GOLD (DES Collaboration, in prep a), for further explanation see section 4.2.1

Y3A2 GOLD column	value	description
`FLAGS_FOOTPRINT`	1	restricts catalog to fiducial survey footprint
`FLAGS_FOREGROUND`	0	excludes regions masked due to foreground objects
`bitand(FLAGS_GOLD, 122)`	0	photometric processing failure exclusion based on SOF
`EXTENDED_CLASS_SOF`	3	high purity galaxy sample based on SOF model

Table 4.2: Y3A2 GOLD catalog query cuts used in obtaining the survey data from the DES Data Management System (DESDM, Mohr et al., 2008)

cover their central richness bins, and enforce a narrower redshift selection to reduce the smearing of observed photometric features (e.g. red sequence) due to mixing of different redshift cluster members. While this smearing is not a limitation for the presented emulator algorithm to model the observed feature distributions, reduced smearing and redshift mixing enables powerful sanity checks in evaluating performance. The binning scheme of this study and of chapter 2 is shown together with the underlying distribution of clusters in Figure 4.2.

In the following we outline the approach used for measuring the distribution of photometric features as a function of the projected separation R from the targeted clusters in a given bin. This calculation corresponds to the step `RadialCorrelator` indicated by green rectangles on the algorithm outline flowchart (Figure 4.4).

4.3.1.1 Database Query and Filtering

The base dataset for this measurement is a subset of the Y3A2 GOLD photometric catalog selected via the flags listed in Table 4.2, queried from the DES Data Management system (DESDM, Mohr et al., 2008). The flags are chosen to yield a high-completeness galaxy sample while excluding photometry failures. For each cluster in a given cluster selection we select all entries from this base catalog which are approximately within a pre-defined search radius $\theta_{query} \approx 6\,\mathrm{deg}$ around the cluster.

In practice this query is implemented by first constructing a low resolution equal-area pixelization of the sky using the HEALPix algorithm (Górski et al., 2005) with `nside=16`, assigning a pixel ID for each entry of the base catalog, and then performing an inclusive disk query around the cluster position to find objects residing in healpix pixels which fully or partially overlap with the search region of radius θ_{query}. This query step drastically reduces the size of the galaxy catalog, and within the selected subset we directly calculate the projected radial separation of each galaxy from the cluster positions.

4.3.1.2 Weighted Catalog Subsampling

Given the very large number of galaxies which are at a large separation from the targeted clusters, directly storing the above dataset is not feasible. Furthermore, the same galaxy can appear multiple times at different radii around different nearby clusters. Therefore, our aim is to construct a

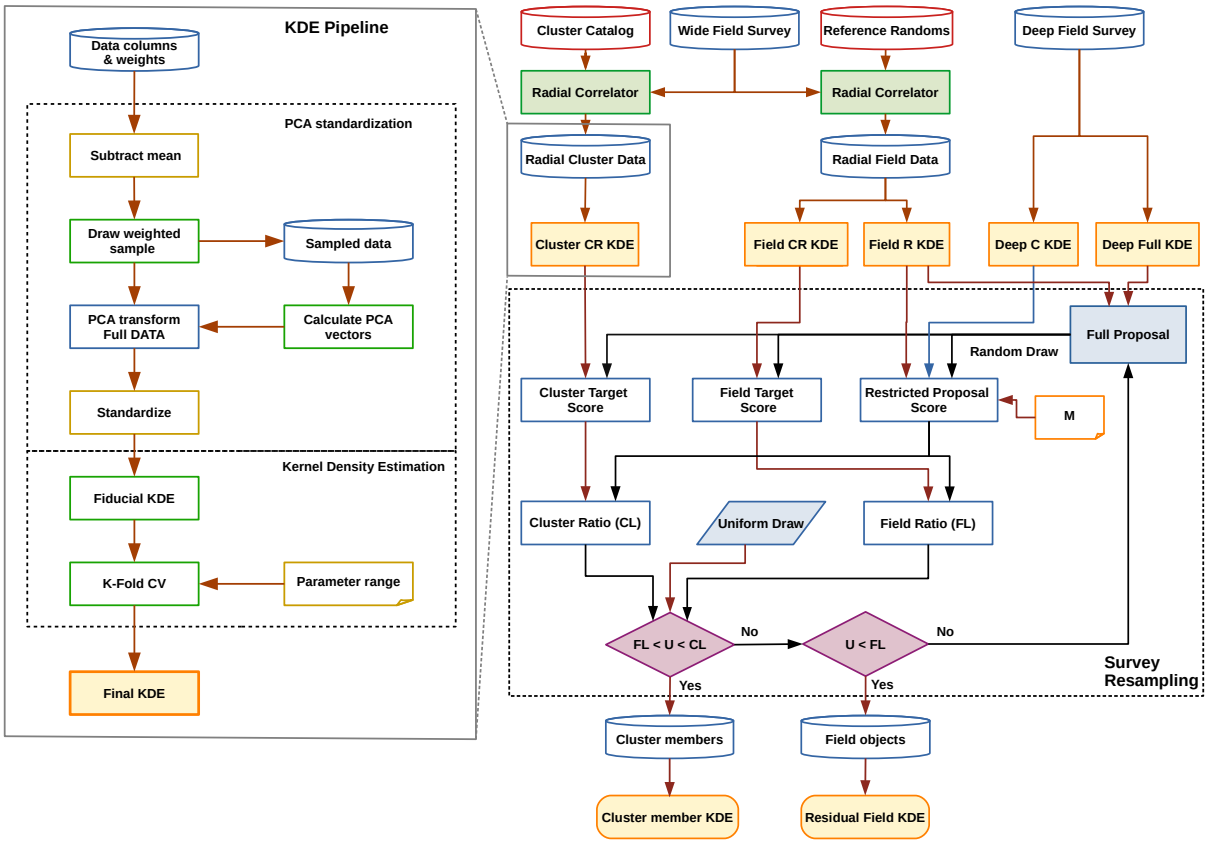

Figure 4.4: Algorithm flowchart of the cluster emulator pipeline. The inputs of the pipeline are the cluster and reference random point catalog (section 4.2.2), the DES wide field (section 4.2.1) and deep field survey (section 4.2.3). The `Radial Correlator` module is described in section 4.3.1, the emulated features are listed in Table 4.1, the `KDE Pipeline` module is described in section 4.3.2, while the `Survey Resampling` module along with the overview of the algorithm is presented in section 4.3.4.

much smaller subsample of the above dataset whose points can be used to reconstruct the same distribution and can be stored efficiently, while also accounting for galaxy multiplicities.

To achieve the above, first we measure the total radial number profile of galaxies which is the sum of the number profiles measured around individual clusters in the selection in radial bins arranged as $R \in [0; 10^{-3})$ arcmin, $[10^{-3}; 0.1)$ arcmin, and in 50 consecutive logarithmically spaced radial bins between 0.1 arcmin and 100 arcmin. Following the above, from each radial range we draw $N_{\rm draw} = \min(N_{\rm bin}; N_{\rm th})$ galaxies where $N_{\rm bin}$ is the number of galaxies in the radial bin, and $N_{\rm th} = 10000$ is a threshold number which allows that the subsampled dataset can be efficiently stored.

The random draw of galaxies is equally partitioned across the $N_{\rm clust}$ clusters in the cluster selection. That is, from the vicinity of each cluster approximately $N_{\rm draw}/N_{\rm clust}$ galaxies are drawn without replacement from each radial bin. To account for the number threshold $N_{\rm th}$, for each drawn galaxy a weight

$$w_{\rm bin} = N_{\rm bin}/N_{\rm draw} \tag{4.1}$$

is associated. This way the subsampled distribution retains its shape, and the number of tracers representing the distribution is reduced in an adaptive way: At low radii all galaxies around clusters can be admitted to the subsample with unity weight, while at large radii only a small fraction is admitted with a correspondingly larger weight. For each selected galaxy the full catalog row is transferred from the GOLD catalog, and through the random draw the same galaxy can enter multiple times, but at different radii.

The outcome of the above algorithm is a galaxy photometry catalog also containing the projected radius R of each entry measured from the targeted cluster sample with a weight of each entry. This dataset is denoted as the `RadialClusterData` on the algorithm flowchart (Figure 4.4), and it is a representation of the radial photometric feature and galaxy density distribution around the targeted clusters. The measurement is repeated also with the sample of reference random points selected in the same richness and redshift range as the cluster sample, yielding the dataset denoted as `RadialFieldData`. This second dataset is representative of the typical *field* galaxy distributions, however through the spatial and redshift distribution of the reference random points it also directly measures and incorporates the impact of survey inhomogeneities and masking.

4.3.2 Kernel Density Representation of Survey Data

Our aim is to generalize the features of a finite set of observed galaxies into an estimate on the multivariate feature probability distribution function (p.d.f.) which they are drawn from. We achieve this task via kernel density estimation, which is preceded by a data standardization step, and the resulting density estimates are validated by a cross validation scheme. These steps correspond to the flowchart outtake on the left side of Figure 4.4 denoted as `KDE Pipeline`.

4.3.2.1 Kernel Density Estimates

Kernel density estimation (KDE) is a type of unsupervised learning algorithm (Parzen, 1962; Hastie et al., 2001). The essence of this algorithm is to estimate and recover the (possibly mul-

tivariate) p.d.f. $p(x)$ based on a set of independent identically distributed random variables (r.v.) $\hat{X} = X_1, X_2, ... X_i$ which are draws from $p(x)$. This recovery is achieved by convolving the discrete set of points with a *Kernel function* $K_h(r)$, where the reconstructed p.d.f. is expressed as

$$p_n(x) = \left(\hat{X} \star K_h\right)(x) = \frac{1}{n}\sum_{j=1}^{n} K_h\left(x - X_j\right) \ . \tag{4.2}$$

Here, h is the bandwidth which sets the smoothing scale during the p.d.f. reconstruction. Parzen (1962) showed that this above form yields a *consistent* estimator of the p.d.f for a wide range of kernel choices. That is, it is asymptotically unbiased and the estimator variance approaches zero as the sample size increases: $|p(x) - p_n(x)| \to 0$, $Var(p_n) \to 0 \ : \ n \to \infty$, where $h \to 0$ in the same limit.

A common choice of kernel function is a Gaussian kernel, which we adopt in the present study:

$$K(r,\, h) = \frac{1}{\sqrt{(2\pi\, h)^d}}\, \exp\left(-\frac{1}{2}\frac{r^2}{h^d}\right), \tag{4.3}$$

where the multivariate nature of the dataset is made explicit by formulating the kernel for d dimensional data with a single bandwidth h, proportional to the variance of the Gaussian. This kernel is continuous, therefore from a set of discrete data points we arrive to a continuous estimate of the p.d.f. Due to the wings of the Gaussian distribution, gaps and undersampled regions are modeled to have non-zero probability, therefore we expect this method to be less sensitive to sparse data compared to alternative kernels with finite support. Finally, the use of a Gaussian kernel enables the efficient drawing of random samples from the reconstructed p.d.f $p_n(x)$, as each draw is translated into a random draw from one of the kernels centered on a sample point, and this is numerically feasible in d dimensions.

For the practical calculation of KDEs we make use of the `scikit-learn` implementation of the above algorithm[1] which is available as a `phyton` package. This implementation combines the KDE algorithm with a decision tree, to optimize the numerical cost of constructing the density estimate for very large multi-dimensional datasets.

4.3.2.2 Data Standardization

The dataset used here for constructing the KDE may have features with very disparate scales. In the case of a photometric dataset this is apparent as the value range and distribution of galaxy magnitudes and galaxy colors is markedly different. This means that any single bandwidth h (smoothing scale) is not equally applicable for all dimensions.

To address this we transform the input features before the KDE step into a set of new features which are better described by a single bandwidth parameter, a process called *standardization*. First we subtract the mean of each feature, then perform a principle component analysis (PCA) to find the eigendirections of the input features (Hastie et al., 2001). Using these eigendirections

[1] `https://scikit-learn.org/stable/modules/density.html`

we then map the features of each galaxy into a set of eigenfeatures. Finally, these are standardized by dividing each eigenfeature by its estimated standard deviation among the sample.

The principle component analysis is performed via the `scikit-learn` implementation[2] which defines it via the singular value decomposition of the dataset:

$$\boldsymbol{X} = \boldsymbol{U} \cdot \boldsymbol{D} \cdot \boldsymbol{V}^T \,. \tag{4.4}$$

Here $\boldsymbol{X}$ is the $n \times d$ matrix composed of the stack of the n observed d dimensional data points $\hat{X}$. $\boldsymbol{U}$ and $\boldsymbol{V}$ are the left- and right handed singular vectors respectively, and the diagonal matrix $\boldsymbol{D}$ contains the singular values in decreasing order. In this formalism the principle component vectors are the column vectors of $\boldsymbol{U} \cdot \boldsymbol{D}$ (Hastie et al., 2001). This formalism does not accommodate for weights in the measured dataset $\hat{X}$, such as the ones defined by Equation 4.1. To remedy this, instead of using the full dataset with weights, we perform a weighted draw of $\sim 10^5$ galaxies with replacement to obtain a representative subset of galaxies. This representative subset is used in the PCA and the principle components of the subset are then used to transform the full dataset containing the weights.

The transformation outlined above corresponds to the `PCA Standardization` section of the algorithm flowchart shown in Figure 4.4.

4.3.2.3 Bandwidth Selection

The performance of a kernel density estimate according to Equation 4.2 depends on the bandwidth h, which in the case of a Gaussian kernel corresponds to the variance of the kernel. While for $n \to \infty$ the optimal bandwidth $h \to 0$ approaches zero (Parzen, 1962), for a finite dataset drawn from a p.d.f of general form (e.g. not a Gaussian), the optimal bandwidth cannot be calculated in an analytic way.

For this reason we perform a data driven optimization to find the optimal bandwidth for the KDE. During this, the starting dataset is split into *training* and *test* data, a KDE $p_n$is constructed via the training data, and a performance score is calculated on the test data x_j:

$$\mathcal{S} = \frac{1}{N} \sum_{j}^{N} \ln p_n(x_j,\, h) \,. \tag{4.5}$$

During the bandwidth selection the score $\mathcal{S}$ is maximized over a grid of proposed bandwidths, which is equivalent to maximizing the probability of the test data given the training data.

Splitting the data into one training and one test set may introduce over-fitting into the optimization, as the algorithm and bandwidth may be artifically tuned to provide good results for a single particular test scenario. A way around this is to adopt a leave-one-out cross-validation strategy (Hastie et al., 2001), where the same base data is split into k equal parts, and from these each part is once considered as the test data, and the remainder is used as the training data. In

[2] `https://scikit-learn.org/stable/modules/decomposition.html`

this approach the score $\mathcal{S}$ is calculated k times on a different training and test combination, and a joint cross-validation score is estimated as

$$CV_k(p_n,\, h) = \frac{1}{k}\sum_{i=1}^{k} \mathcal{S}_i\,. \tag{4.6}$$

Here S_i is evaluated on points in the i-th subset, while the KDE is trained on points not in the i-th subset. The final KDE is then constructed from the full dataset using the bandwidth maximizing the CV_k score, where we adopt $k = 5$.

Originating from the PCA standardization step, bandwidths can be expressed relative to the standard deviation $\sigma = 1$ of the various standardized eigenfeatures. Based on this we evaluate the cross validation score on a logarithmically spaced bandwidth grid from 0.01σ to 1.2σ for each KDE shown on Figure 4.4. We find that $h \approx 0.1\sigma$ simultaneously provides a good bandwidth estimate for the deep field and the wide field KDEs, hence we adopt it as a global bandwidth for further calculations.

This above cross-validation bandwidth selection step, along with the construction of the final KDE using the optimized bandwidth corresponds to the `Kernel Density Estimation` section of the algorithm flowchart shown on Figure 4.4.

4.3.3 Cluster and Field Population Estimates

The aim of this section is to model the radial feature distribution of *cluster member* galaxies for different samples of galaxy clusters, and to separate this from the distribution of foreground and background galaxies which we expect to be be similar to *field* galaxies of the median survey line of sight. The input data product for the following calculations is the feature p.d.f. estimated from the various deep field and wide field galaxy catalogs using the kernel density estimate pipeline introduced in section 4.3.2 The full list of features along with their definitions from DES Y3 data products is shown in Table 4.1.

First, in section 4.3.3.1, we formally define the statistical background subtraction used to estimate the feature distribution of cluster members. This formal definition is not yet straightforward to evaluate in practice, and has to be updated to a more practical approach. Meanwhile, in section 4.3.3.2 we introduce the extrapolation scheme used in combining deep field and wide field data. Finally in section 4.3.3.3 we introduce and overview a rejection sampling approach which can be used to combine and jointly perform the operations defined in section 4.3.3.1 and section 4.3.3.2.

4.3.3.1 Statistical Background Subtraction

When measuring the correlation of galaxy positions and features with cluster positions as performed in section 4.3.1, all line-of-sight galaxies are included. Hence, the cluster members are mixed with foreground and background galaxies, which dilute and contaminate the measurement. Photometric redshift estimates available for the wide field DES surveys (Y1 Hoyle et al., 2018 or Y3 DES Collaboration, in prep c) are not precise enough to isolate a sufficiently pure and

label	p.d.f.	survey	feature	range	comment
`Cluster CR KDE`	$P(\vec{c}_{\rm wide}, R \,\vert\, \lambda, z)$	wide field		$i < 22.5$	measured around clusters with λ and z
			$\vec{c}_{\rm wide}$	$[-1; 3.5]$	reference colors for redshift range from Table 4.4
			R	$[10^{-1.5}; 10^{1.2}]$	projected radius from cluster center [arcmin]
`Field CR KDE`	$P(\vec{c}_{\rm wide}, R \,\vert\, \lambda, z)$	wide field		$i < 22.5$	measured around reference random points with λ and z
			$\vec{c}_{\rm wide}$	$[-1; 3.5]$	reference colors for redshift range from Table 4.4
			R	$[10^{-1.5}; 10^{1.2}]$	projected radius from reference point [arcmin]
`Field R KDE`	$P(R \,\vert\, \lambda, z)$	wide field		$i < 22.5$	measured around reference random points with λ and z
			R	$[10^{-1.5}; 10^{1.2}]$	projected radius from reference point [arcmin]
`Deep Full KDE`	$P(m, \vec{c}, \vec{s}, z_g)$	deep field		i < 25	
			m	$[17; 25]$	i-band magnitude
			$\vec{c}$	$[-1; 3.5]$	$g-r$, $r-i$, $i-z$ colors
			$\vec{s}_1$	$[0; 1]$	absolute ellipticity
			$\vec{s}_2$	$[0; 1]$	bulge disk fraction
			$\vec{s}_3$	$[-1; 5]$	size (area) proxy
			z_g	$[0; 3.5]$	Monte Carlo draw from galaxy redshift $p(z)$
`Deep C KDE`	$P(c_{\vec{\rm wide}})$	deep field		i < 22.5	
			$\vec{c}_{\rm wide}$	$[-1; 3.5]$	reference colors for redshift range from Table 4.4

Table 4.3: Kernel density estimates and the covered features and value ranges labeled as they appear in Figure 4.4

redshift bin	cluster redshift range	components	comment
1	$z \in [0.3;\ 0.35)$	$\vec{c}_{\mathrm{wide};z1} = (c_1,\ c_2)$	$g-r$, $r-i$ colors
2	$z \in [0.45;\ 0.5)$	$\vec{c}_{\mathrm{wide};z2} = (c_1,\ c_2)$	$g-r$, $r-i$ colors
3	$z \in [0.6;\ 0.65)$	$\vec{c}_{\mathrm{wide};z3} = (c_2,\ c_3)$	$r-i$, $i-z$ colors

Table 4.4: Reference colors for the redshift ranges of the different cluster selections

complete sample of cluster member galaxies across the full range of galaxy populations (e.g. not only the red sequence). Therefore, to avoid the above limitation, we perform a *statistical background subtraction* (Hansen et al., 2009). That is, instead of separating the line of sight galaxy population into disjoint sets of cluster and non-cluster galaxies, we only aim to estimate the overall feature *distribution* of cluster member galaxies, while allowing each galaxy to contribute to both distributions with non-zero weight.

We formulate the statistical background subtraction as performed on the p.d.f.-s of different galaxy feature distributions $p(\vec{\theta},\ R \mid \lambda,\ z)$, where $\vec{\theta}$ is the list of features considered, and R is the projected separation from the targeted positions on the sky, while λ and z are the richness and redshift of the targeted clusters. In practice, each p.d.f. corresponds to a KDE constructed on the relevant features and radius following section 4.3.2, after estimating the radial distributions according to section 4.3.1. In this framework we describe the line-of-sight galaxy population as a two-component system, with a cluster member population, and a field population. The sum of these two (after normalization) yields the total galaxy population $p_{clust}(\vec{\theta},\ R \mid \lambda,\ z)$ measured in projection around clusters. Subtracting an appropriate estimate of the field population from the total population thus then yield an estimate on the cluster member population's $p_{memb}(\vec{\theta},\ R \mid \lambda,\ z)$. Following the approach already introduced in chapter 3 we approximate the field population as the projected feature distribution around a corresponding sample of reference random points $p_{rand}(\vec{\theta},\ R \mid \lambda,\ z)$. In this study, both the cluster positions and the corresponding set of reference random points are provided by the redMaPPer algorithm. Random points are also labeled by richness and redshift to indicate which cluster selection they are representative of.

Based on the above, the feature distribution of cluster member galaxies can be written as the excess probability of finding a galaxy around a cluster compared to a random point:

$$p_{\mathrm{memb}}(\vec{\theta},\ R) = \frac{\hat{n}_r}{\hat{n}_c - \hat{n}_r}\left[\frac{\hat{n}_c}{\hat{n}_r} p_{\mathrm{clust}}(\vec{\theta},\ R) - p_{\mathrm{rand}}(\vec{\theta},\ R)\right] \tag{4.7}$$

where in practice both p.d.f-s on the right hand side are measured in the wide field dataset according section 4.3.2, and $\hat{n}_c$ and $\hat{n}_r$ refer to the mean number of galaxies detected within R_{max} around clusters and random points. These factors are necessary as p.d.f.-s are normalized to unity, but clusters are overdense compared to median line-of-sight (i.e. $\hat{n}_c > \hat{n}_r$), which has to be accounted for during the background subtraction. Equation 4.7 is formulated as an operation on the p.d.f., but it is mathematically equivalent to directly subtracting the multi-dimensional histograms of the two distributions. For practical application the p.d.f. formalism carries benefits compared to histograms: The kernel density estimate of the p.d.f. is less susceptible to shot-noise, and the

representation of the distributions is less impacted by the curse of dimensionality as there is no need to uniformly cover the parameter space with a histogram grid (Hastie et al., 2001).[3]

While as described above the operation defined in Equation 4.7 is correct, yet it is impractical to carry out numerically because of the way the KDE representations of p.d.f-s are constructed and stored. For this reason we will define an equivalent mathematical operation in a later section (section 4.3.3.3) which is used in the actual calculations. However, before introducing that, we first investigate the additional task of survey depth extrapolation which we also have to perform in modeling the cluster member galaxy feature distributions.

4.3.3.2 Survey Depth and Feature Extrapolation

The statistical background subtraction formally defined in Equation 4.7 estimates the the feature distribution of the cluster member population $p_{\rm memb}(\vec{\theta},\, R \,|\, \lambda,\, z)$, under the assumption that the ingredient p.d.f.s $p_{\rm clust}(\vec{\theta},\, R)$ and $p_{\rm rand}(\vec{\theta},\, R)$ are calculated correctly from the kernel density estimation step outlined in section 4.3.2. However the approach carries the limitation that it is only applicable for those features $\vec{\theta}$ and their respective value ranges which are covered by the wide field dataset. Furthermore, the formalism implicitly assumes that the p.d.f-s are dominated by the intrinsic distribution of properties, and not by measurement errors. To fulfill this requirement the wide field data must be restricted to a parameter range where photometry errors play a subdominant role, and the completeness of the survey is high.

The above requirements necessitate excluding the bulk of the galaxy population from the background subtraction scheme, as for the DES Y3 GOLD wide field survey the 99% completeness depth is $i < 22.5$ (DES Collaboration, in prep a), and this includes only a minor fraction of all galaxies in any line of sight. Especially important in relation to this study are galaxies whose flux is great enough to meaningfully contribute to the total light in a part of the sky, yet they are not fully resolved or they cannot be detected with confidence using standard survey photometry pipelines (Suchyta et al., 2016). Nevertheless these partial or non-detections have a significant impact on the photometric performance of survey data products (Hoekstra et al., 2017; Euclid Collaboration et al., 2019; Eckert et al., 2020). Therefore they must be modeled and included in the statistical description of a line-of-sight. A distinct undetected population of galaxies is associated with galaxy clusters, which are the faint-end of the cluster member galaxy population. The feature distribution of these galaxies is markedly different from the distribution of faint galaxies in the *field* (cosmic median) line-of-sight. Therefore to model cluster lines-of-sight both populations have to be accounted for.

To characterize the properties of galaxies faint enough to only have incomplete detections in the DES wide field survey, we make use of the DES deep field survey. Owing to ≈ 80 greater exposure time over many epochs, the completeness depth of the deep fields in the `COADD_TRUTH` mode is ~ 2 mag deeper than the wide fields (DES Collaboration, in prep b), and the measured fluxes and models of galaxy morphology are less impacted by noise at fixed magnitude compared to the DES Y3 GOLD wide field catalog.

[3]The memory required for representing a p.d.f. based on N points in d dimensions scales as: $\propto n_{bin}^{d}$ for a histogram with uniform bin sizes, while for the KDE method scales as $\propto N$. In case of a decision tree based KDE as used by `scikit-learn` the scaling goes with $\propto n_{leaf}$ where $n_{leaf} \leq N$

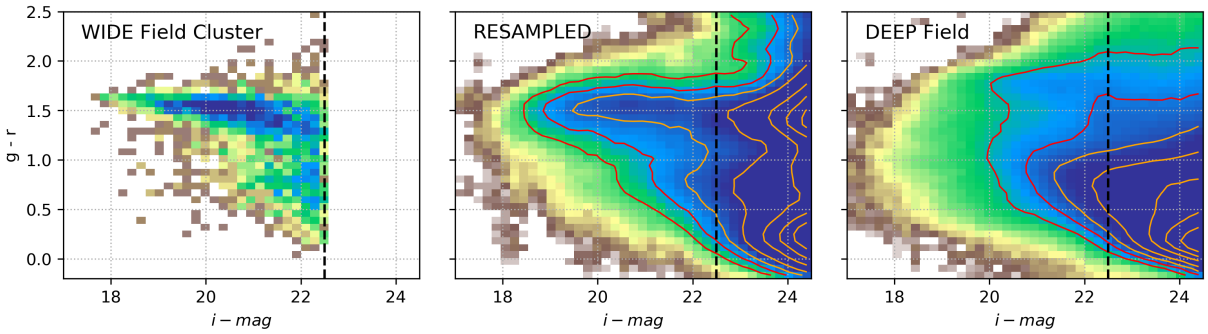

Figure 4.5: An illustration of the re-weighting approach according to Equation 4.17. *left:* color-magnitude diagram of galaxies measured in projection with $R \in [10^{-0.5}; 1)$ arcmin around redMaPPer galaxy clusters with $\lambda \in [45; 60$ and $z \in [0.3; 0.35)$ in the DES wide field survey. *center:* transformed deep field distribution according to Equation 4.17. *right:* color magnitude diagram of galaxies measured in the DES deep field survey. *dashed:* wide field completeness magnitude $i \approx 22.5$. The color scale and contour levels are identical in the three panels. It is visible, that for the $i < 22.5$ magnitude range, the color based re-weighting was able to reproduce the color-magnitude distribution of the line-of-sight

Our aim is therefore to combine the galaxy distributions of the deep field and the wide field in a way where we obtain information about the extrapolation of the various feature distributions $p_{\mathrm{memb}}(\vec{\theta}, R \,|\, \lambda, z)$, $p_{\mathrm{clust}}(\vec{\theta}, R \,|\, \lambda, z)$ and $p_{\mathrm{rand}}(\vec{\theta}, R \,|\, \lambda, z)$ to fainter magnitudes. Furthermore, even for $i < 22.5$ there are features such as the `ngmix` SOF/MOF morphology model which are measured more robustly for deep fields, and we aim to make use of this additional information.

To achieve this task, we first denote the target of the operation $\tilde{p}_D(\vec{\theta}, R \,|\, \lambda, z)$ where the subscript D indicates that the distribution is estimated down to a completeness limit of $i \approx 24.5$. In the following we decompose $\vec{\theta}$ into two sets of features: $\vec{\theta}_{\mathrm{wide}}$ which can be measured from the wide field dataset, and $\vec{\theta}_{\mathrm{deep}}$ which can only be reliably measured from the deep field dataset, thus

$$\tilde{p}_D(\vec{\theta}, R \,|\, \lambda, z) \equiv \tilde{p}_D(\vec{\theta}_{\mathrm{deep}}, \vec{\theta}_{\mathrm{wide}}, R \,|\, \lambda, z)\,. \tag{4.8}$$

Here we note that R, λ, and z are features and quantities which also only originate from the wide field dataset. We note that all features in θ_{wide} can also be measured with confidence in the deep field, but the reverse is not necessarily true.

We formulate Equation 4.8 as a transformation of naive proposal distribution, expressed by the factorization:

$$\begin{aligned}\tilde{p}_D(\vec{\theta}_{\mathrm{deep}}, \vec{\theta}_{\mathrm{wide}}, R \,|\, \lambda, z) = {} & p_{D:prop}(\vec{\theta}_{\mathrm{deep}}, \vec{\theta}_{\mathrm{wide}}, R \,|\, \lambda, z) \\ & \times \tilde{F}(\vec{\theta}_{\mathrm{deep}}, \vec{\theta}_{\mathrm{wide}}, R \,|\, \lambda, z)\,.\end{aligned} \tag{4.9}$$

Here we separate the task into two parts, where the proposal distribution $p_{D:\mathrm{prop}}$ carries information measured from the deep fields, and the multiplicative term $\tilde{F}$ represents the required transformation of the p.d.f. to fulfill the previously defined criterions.

As there is no cluster information from the deep field survey, the proposal p.d.f. cannot depend on λ and z, that is

$$p_{D;\mathrm{prop}}(\vec{\theta}_{\mathrm{deep}},\, \vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z) = p_{D;\mathrm{prop}}(\vec{\theta}_{\mathrm{deep}},\, \vec{\theta}_{\mathrm{wide}},\, R)\,, \tag{4.10}$$

and for the same reason in the proposal distribution $\vec{\theta}_{\mathrm{deep}}$ and $\vec{\theta}_{\mathrm{wide}}$ cannot be correlated with R:

$$p_{D:\mathrm{prop}}(\vec{\theta}_{\mathrm{deep}},\, \vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z) = p_D(\vec{\theta}_{\mathrm{deep}},\, \vec{\theta}_{\mathrm{wide}}) \cdot p_{D:\mathrm{prop}}(R)\,. \tag{4.11}$$

Here $p_D(\vec{\theta}_{\mathrm{deep}},\, \vec{\theta}_{\mathrm{wide}})$ can be directly measured from the deep field survey, and $p_{D:\mathrm{prop}}(R)$ is chosen to capture the approximately uniform surface density of galaxies, e.g. $p_{D:\mathrm{prop}}(R) \propto R$.

The remaining task is to find an appropriate multiplicative term $\tilde{F}(\vec{\theta}_{\mathrm{deep}},\, \vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z)$ which transforms the proposal distribution $p_{D:\mathrm{prop}}$ into the target distribution $\tilde{p}_D$. Since $\tilde{p}_D$ depends on λ, z and R, and $p_{D:\mathrm{prop}}$ is independent of these, the $\tilde{F}$ term must contain all such information. Furthermore, the correlation between $\vec{\theta}_{\mathrm{deep}}$ and R cannot be measured from wide field data, therefore we approximate $\tilde{F}$ as

$$\tilde{F}(\vec{\theta}_{\mathrm{deep}},\, \vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z) \approx F(\vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z)\,. \tag{4.12}$$

A necessary consistency constraint placed on F is expressed as

$$\tilde{p}_D(\vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda\, z)\Big|_W = p_{D;\mathrm{prop}}(\vec{\theta}_{\mathrm{wide}},\, R) \times F(\vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z)\Big|_W \tag{4.13}$$

$$= p_W(\vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z) \tag{4.14}$$

where the W subscript indicates a p.d.f. estimated from wide field data, and the $|_W$ subscript denotes that the otherwise greater magnitude range is restricted to the wide field completeness magnitude of $i \approx 22.5$. From the above constraint it is then possible to find the simplest form of F, as

$$F(\vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z) = \frac{1}{\hat{V}} \frac{p_W(\vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z)}{p_{D;\mathrm{prop}}(\vec{\theta}_{\mathrm{wide}},\, R)\Big|_W} \tag{4.15}$$

$$= \frac{1}{\hat{V}} \frac{p_W(\vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z)}{p_D(\vec{\theta}_{\mathrm{wide}})\Big|_W \cdot\, p_{D;\mathrm{prop}}(R)} \tag{4.16}$$

where $\hat{V}$ is a normalization factor to account for the different volumes of the wide field and deep field parameter spaces, e.g. the difference in the limiting depth of $i < 22.5$ versus $i < 24.5$.

From the combination of Equation 4.11 and Equation 4.16 we can then write our estimate on the target distribution as

$$\tilde{p}_D(\vec{\theta}_{\mathrm{deep}},\, \vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z) \approx \frac{p_D(\vec{\theta}_{\mathrm{deep}},\, \vec{\theta}_{\mathrm{wide}})\, p_W(\vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z)}{\hat{V} \cdot\, p_D(\vec{\theta}_{\mathrm{wide}})\Big|_W}\,, \tag{4.17}$$

where $p_{D;\mathrm{prop}}(R)$ drops out, and the approximation is composed entirely of p.d.f-s which can be directly measured from the wide field or deep field data.

Equation 4.17 is only the simplest of the allowed extrapolation equations, and with external information on the correlations of the various features more complex formulations are possible. Nevertheless this form already possesses a set of useful qualities, which allow us to incorporate it into our analysis and calculations in the way described below. We note however that this is not a unique, albeit a practically useful way of doing so.

First, as the above Equation 4.17 distinguishes wide field and deep field features, we define what is meant by this in practice. Here we take $\vec{\theta}_{\mathrm{wide}} = \vec{c}_{\mathrm{wide}}$ as a set of colors measured in both the wide field survey and deep field survey, and $\vec{\theta}_{\mathrm{deep}} = (m, \vec{s}, \vec{c}_{\mathrm{deep}}, z_g)$ is a vector composed of magnitudes, colors, morphology parameters and redshifts measured in the deep field survey according to Table 4.1. From the wide field observations we have information about the distribution of galaxy colors around clusters as function of projected radius R within a cluster sample, but only down to the wide field completeness magnitude $i < 22.5$. For the deep field, we have information about the full set of galaxy photometry and morphology features, but no information relating to clusters (that is no R, λ or z), but down to a deeper completeness magnitude of $i < 24.5$. The meaning of Equation 4.17 is to combine these two sets of information. As is visible in Equation 4.10 this can be thought of as a re-weighting of the deep field feature p.d.f. by a weight factor, where the weight factor captures the cluster related information. In the present formulation, this re-weighting is done based on the $\vec{c}_{\mathrm{wide}}$ set of colors, which are measured in both the wide and the deep field data (for the different cluster redshift samples these are listed in Table 4.4). This framework conserves the color dependent luminosity function, and in general it obeys that

$$\tilde{p}_D(\vec{\theta}_{\mathrm{deep}}|\,\vec{\theta}_{\mathrm{wide}},\, R,\, \lambda,\, z) \equiv p_D(\vec{\theta}_{\mathrm{deep}}|\,\vec{\theta}_{\mathrm{wide}}) \tag{4.18}$$

Ṡince magnitudes are part of $\vec{\theta}_{\mathrm{deep}}$, this means that the the final p.d.f. estimate inherits the luminosity function of the deep fields, along with all additional features which are measured in the deep fields. Hence the survey depth is extrapolated in a smooth and continuous way.

An illustration of this approach is shown on Figure 4.5, where the left panel shows the target distribution, the color-magnitude diagram of galaxies measured in projection with $R \in [10^{-0.5};\, 1)$ arcmin around redMaPPergalaxy clusters with $\lambda \in [45;\, 60$ and $z \in [0.3;\, 0.35)$ in the DES wide field survey. The right panel of this figure shows the comparable color magnitude diagram of galaxies measured in the DES deep field survey, with the wide field completeness magnitude shown as the vertical dashed line. The center panel of this figure shows the transformed deep field distribution according to Equation 4.17, where the radial color distribution around the cluster sample was used as the target p.d.f. The color scale is identical in the three panels, and iso-probability contours are added to trace the parts of the distribution where the color scale saturates. It can be seen that for the $i < 22.5$ magnitude range, the color based re-weighting is able to reproduce the color-magnitude distribution of the line-of-sight: reconstructing the red-sequence and blue cloud of the cluster line of sight out of deep field galaxies. Furthermore, once can see that by applying this re-weighting to the deep field galaxy distribution, the observed magnitude distribution of each color cell is retained down to the deep field cutoff magnitude.

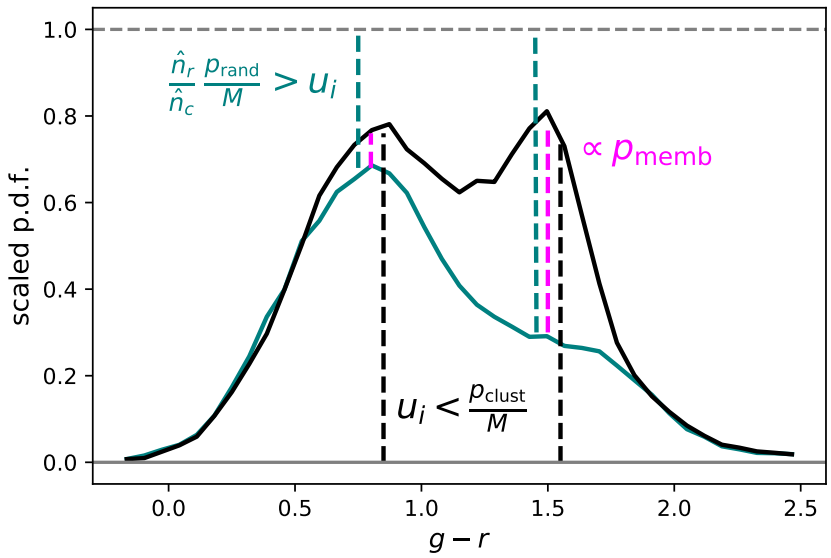

Figure 4.6: Illustration of statistical background subtraction via rejection sampling according to Equation 4.22. *(black:)* scaled p.d.f measured around clusters $p_{\rm clust}/M$. *(teal:)* scaled p.d.f. measured around reference random points $p_{\rm rand}/\frac{\hat{n}_c}{\hat{n}_r}M$. The black and teal vertical dashed lines indicate the probability of the different events needed independently fulfill the two inequalities in the r.h.s of Equation 4.22, while the length of the magenta dashed line is proportional to joint probability of the two events.

4.3.3.3 Rejection Sampling

The aim of this section is to jointly perform the statistical background subtraction (section 4.3.3.1) and the survey depth extrapolation (section 4.3.3.2), generating random samples from the thus isolated and extrapolated cluster member feature distribution.

Generating a multivariate random variable which is distributed according to a p.d.f. of generic form is not trivial (MacKay, 2002). The KDE representation of p.d.f.-s adopted in section 4.3.2 avoids this limitation by reducing the random draw from the p.d.f. into a random draw from a Gaussian kernel. However the KDE approach introduces an independent difficulty, as the p.d.f. is not known analytically. Evaluating it is computationally much more expensive than drawing random samples, as it requires summing over the contribution of all data points whose kernel has non-zero contribution. Furthermore, the KDE representation does not lend itself readily to performing abstract operations on the p.d.f. such as subtracting or dividing two distinct KDEs, as they are built on different datasets and decision trees.

In light of the above, we adopt an approach where instead of directly performing the background subtraction operation defined by Equation 4.7 or the survey extrapolation defined by Equation 4.17, we aim to generate random samples from the target distribution $\tilde{p}_{D;\,\rm memb}$ which is the estimate on the feature distribution of the cluster member galaxies extrapolated with deep field features and limiting magnitudes. For this we make use of an approach known as *rejection sampling* or alternatively *accept - reject sampling* or *von Neumann rejection* (MacKay, 2002).

The essence of rejection sampling is to generate random variables distributed according to a *target distribution* $p_{\rm targ}(\vec{\alpha})$ by performing random draws from a *proposal distribution* $p_{\rm prop}(\vec{\alpha})$. The crucial benefit is that this approach only requires that the target probability can be numerically evaluated, and it does not rely on knowing its analytic form.

For each random realization $\vec{\alpha}_i$ drawn from $p_{\rm prop}(\vec{\alpha})$, an acceptance criterion is defined using a second draw u_i from a uniform distribution $\mathcal{U}[0;\,1)$. Proposal draws fulfilling the inequality

$$u_i < \frac{p_{\rm targ}(\vec{\alpha}_i)}{\mathcal{M}\cdot p_{\rm prop}(\vec{\alpha}_i)} \tag{4.19}$$

are accepted, while draws which do not fulfill it are discarded. The $\vec{\alpha}_i$ values meeting the above selection are guaranteed to be distributed according to $p_{\rm targ}$ (MacKay, 2002). The factor $\mathcal{M}$ appearing in the above inequality is defined to ensure that

$$\mathcal{M}\cdot p_{\rm prop}(\vec{\alpha}) > p_{\rm targ}(\vec{\alpha})\ :\ \forall\vec{\alpha}\,. \tag{4.20}$$

There is a simple geometric interpretation of rejection sampling, which is that $\mathcal{M}\cdot p_{\rm prop}$ defines an upper envelope function for $p_{\rm targ}$, and the acceptance rate for $\vec{\alpha}_i$ is proportional to the distance between this envelope function and $p_{\rm targ}$. That is where the upper envelope function closely approximates the target distribution the acceptance rate is close to unity. Naturally this means, that the numerical efficiency of the rejection sampling algorithm greatly benefits from finding a proposal distribution which is similar to the target distribution, although this is not a mathematical requirement of the method.

We now formulate the statistical background subtraction and extrapolation in the language of rejection sampling. As introduced in section 4.3.3.1, the cluster member galaxy population can be statistically defined as the feature dependent galaxy excess compared to a reference random line-of-sight. This can be equivalently expressed as a galaxy with feature $\vec{\beta}_i$ having membership probability

$$p_{\rm memb}(\vec{\beta}_i) = P\left[p_{\rm rand}(\vec{\beta}_i) < u_i\cdot\frac{\hat{n}_c}{\hat{n}_r}\,sup\left(p_{\rm clust}(\vec{\beta}_i)\right) < \frac{\hat{n}_c}{\hat{n}_r}p_{\rm clust}(\vec{\beta}_i)\right] \tag{4.21}$$

where in the right hand side $P[\cdot]$ refers to the probability that the double inequality is fulfilled, $\hat{n}_c$ and $\hat{n}_r$ refer to the normalization factors of $p_{\rm clust}$ and $p_{\rm rand}$ respectively, $sup(\cdot)$ refers to the supremum of the function over the parameter space, and it is assumed that the proposed $\vec{\beta}$ is distributed uniformly over the parameter space. Since in practice $p_{\rm clust}$ is not known exactly, we can rewrite the above equation by replacing it with an appropriately chosen value $\mathcal{M}$ which fulfills that $\frac{\hat{n}_c}{\hat{n}_r}p_{\rm clust} < \mathcal{M}$ and $p_{\rm rand} < \mathcal{M}$:

$$p_{\rm memb}(\vec{\beta}_i) \propto P\left[p_{\rm rand}(\vec{\beta}_i) < u_i\cdot\frac{\hat{n}_c}{\hat{n}_r}\mathcal{M} < \frac{\hat{n}_c}{\hat{n}_r}p_{\rm clust}(\vec{\beta}_i)\right]\,, \tag{4.22}$$

in which case the equality turns into a proportionality. The geometric meaning of this approach is illustrated on Figure 4.6, which shows the appropriately scaled version of the $p_{\rm clust}$ and $p_{\rm rand}$ p.d.f.-s as the black and teal curves respectively. The probability of the two events appearing in

the inequalities in the right hand side (r.h.s.) of Equation 4.22 are indicated by the lengths of the black and the teal lines, while the joint probability of the two events happening at the same time is shown by the length of the magenta line.

The acceptance rate can be increased by using a cleverly choosen proposal distribution p_{prop} which is other than uniform. In this case the above equation needs to be modified as

$$p_{\mathrm{memb}}(\vec{\beta}_i) \propto p\left[p_{\mathrm{rand}}(\vec{\beta}_i) < u_i \cdot \frac{\hat{n}_c}{\hat{n}_r}\mathcal{M} \cdot p_{\mathrm{prop}} < \frac{\hat{n}_c}{\hat{n}_r} p_{\mathrm{clust}}(\vec{\beta}_i)\right] . \tag{4.23}$$

The above form can be transformed to resemble Equation 4.19 and make the similarities with the simple rejection sampling approach manifest:

$$p_{\mathrm{memb}}(\vec{\beta}_i) \propto p\left[\frac{p_{\mathrm{rand}}(\vec{\beta}_i)}{\frac{\hat{n}_c}{\hat{n}_r}\mathcal{M} \cdot p_{\mathrm{prop}}} < u_i < \frac{p_{\mathrm{clust}}(\vec{\beta}_i)}{\mathcal{M} \cdot p_{\mathrm{prop}}}\right] \tag{4.24}$$

where u_i is drawn from $\mathcal{U}[0;\,1)$, and $p_{rand}(\vec{\beta}_i)$ and $p_{clust}(\vec{\beta}_i)$ are the feature p.d.f. of galaxies measured in projection around reference random points and galaxy clusters respectively. $\hat{n}_c/\hat{n}_r$ is the average relative overdensity of galaxy count in the cluster line-of-sight compared to a reference random line of sight.

Equation 4.24 has limited practical use, since both p_{rand} and p_{clust} can be measured directly from data, negating the the need for resampling. Nevertheless Equation 4.24 can be used to incorporate the survey extrapolation according to Equation 4.17. For this we adopt the proposal distribution as defined by Equation 4.11, that is

$$\begin{aligned} p_{\mathrm{prop}} &= p_{\mathrm{prop}}(\vec{\theta}_{\mathrm{deep}},\, \vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z) \\ &= p_D(\vec{\theta}_{\mathrm{deep}},\, \vec{\theta}_{\mathrm{wide}}) \cdot p_{W;\,\mathrm{rand}}(R \,|\, \lambda,\, z) \\ &= p_D(m,\, \vec{c},\, \vec{s},\, z_g) \cdot p_{W;\,\mathrm{rand}}(R \,|\, \lambda,\, z) \end{aligned} \tag{4.25}$$

which we use to draw the proposal random samples from. Furthermore we define a restricted proposal distribution which contains only features contained within $\vec{\theta}_{ref}$, that is

$$\begin{aligned} p_{rp} &= p_{rp}(\vec{\theta}_{\mathrm{wide}},\, R \,|\, \lambda,\, z) \\ &= p_D(\vec{c}_{\mathrm{wide}}) \cdot p_{W;\,\mathrm{wide}}(R \,|\, \lambda,\, z)\,, \end{aligned} \tag{4.26}$$

which can be directly compared with p_{clust} and p_{rand}.

Combining the above, we can generate random samples from the survey extrapolated p_{memb}, by drawing samples $\{m_i,\, \vec{c}_i,\, \vec{s}_i,\, z_{g;i},\, R_i\}$ from Equation 4.25, and considering the subset which fulfills the extrapolated membership criterions

$$\frac{\hat{n}_r}{\hat{n}_c}\frac{p_{W;\,\mathrm{rand}}(\vec{c}_{\mathrm{wide;i}},\, R_i \,|\, \lambda,\, z)}{\mathcal{M} \cdot p_D(\vec{c}_{\mathrm{wide};i}) \cdot p_{W;\,\mathrm{rand}}(R_i \,|\, \lambda,\, z)} < u_i \tag{4.27}$$

and

$$u_i < \frac{p_{W;\,\mathrm{clust}}(\vec{c}_{\mathrm{wide};i},\, R_i \,|\, \lambda,\, z)}{\mathcal{M} \cdot p_D(\vec{c}_{\mathrm{wide};i}) \cdot p_{W;\,\mathrm{rand}}(R_i \,|\, \lambda,\, z)}\,. \tag{4.28}$$

As a null-test, we can also perform the same resampling for the galaxies around random points, which using the same proposal distribution as above, is defined by the criterion

$$u_i < \frac{\hat{n}_r}{\hat{n}_c} \frac{p_{W;\,\mathrm{rand}}(\vec{c}_{\mathrm{wide};i},\, R_i \,|\, \lambda,\, z)}{\mathcal{M} \cdot p_D(\vec{c}_{\mathrm{wide};i}) \cdot p_{W;\,\mathrm{rand}}(R_i \,|\, \lambda,\, z)} \,. \tag{4.29}$$

In the above formalism the factor $\mathcal{M}$ must be chosen appropriately to ensure that Equation 4.20 holds. In practice there is no given recipe for $\mathcal{M}$, and the suitable value must be found for the actual samples proposed. In addition, as the p.d.f.-s in this analysis are estimated numerically, the measurement noise can lead to small fluctuations in them. This, especially in the wings of the distributions where probabilities are small, can manifest as $p_{\mathrm{targ}}/p_{\mathrm{prop}}$ being very large, necessitating a very large value of $\mathcal{M}$. This would lead to an artificially low acceptance rate imposed by noise and poorly sampled regions. To regularize this behaviour we relax the requirement on $\mathcal{M}$ and in practice only require that Equation 4.20 be fulfilled for 99% of the proposed points. We explore the $\mathcal{M}$ range in an iterative fashion up to 500, and find only no significant change in the distribution of the samples for $\mathcal{M} > 40$, hence in the following we adopt $\mathcal{M} \approx 100$ for the criterions Equation 4.27, Equation 4.28 and Equation 4.29.

4.3.4 Algorithm Overview

Based on the methods introduced in the previous sections, we now overview the complete algorithm used to estimate the feature distribution of cluster member galaxies. The flowchart for this algorithm is shown on Figure 4.4.

The primary inputs for the calculation are the photometric catalog of DES Y3 GOLD indicated as `Wide Field Survey` introduced in section 4.2.1, the photometry and photometric redshift catalog of the DES Y3 deep and supernova survey indicated as `Deep Field Survey` on Figure 4.4 introduced in section 4.2.3, and the redMaPPer cluster and reference random point catalog indicated as `Cluster Catalog` and `Reference Randoms` respectively, which are introduced in section 4.2.2.

4.3.4.1 Data Preparation

In the first calculation step, the wide field dataset and the cluster and reference random catalogs are combined by the operation indicated as `Radial Correlator` on Figure 4.4, which is introduced in section 4.3.1. The essence of this step is to construct a set of features for each galaxy in the wide field catalog, and to measure the radial distribution of galaxies around clusters and reference random points. The features for wide field detected galaxies consists of i-band magnitude, $g-r$, $r-i$, $i-z$ colors, and the additional projected radius measured from the reference point. The exact definition of these features from the DES Y3 GOLD catalog is shown in Table 4.1. The outcome of this calculation is two new dataset composed of a catalog of galaxies with the above features defined with the target cluster positions denoted as `Radial Cluster Data` and with the reference random points denoted as `Radial Field Data`. This latter is motivated by the approximation that the environments of the reference random points measure the

galaxy properties of the typical *field* lines-of-sight. For the deep field survey we directly transform the DES deep field data into the set of features listed in Table 4.1. The used features are i-band magnitude, $g-r$, $r-i$, $i-z$ colors, a photometric redshift estimate z_g for each galaxy, and three morphology parameters used in the `ngmix` SOF/MOF galaxy model: absolute ellipticity, area, bulge / disk fraction.

In the following step, from the galaxy feature dataset, kernel density estimates of the underlying feature p.d.f.-s are constructed according to the approach introduced in section 4.3.2. From the three datasets a total of five separate KDEs are constructed as shown on Figure 4.4, which are listed along with the list of features and value ranges in Table 4.3. In short, these KDEs correspond to the estimate on the radial color distribution of galaxies around selections of redmapper clusters in the DES wide field survey $P(\vec{c}_{ref}, R \,|\, \lambda, z)$ denoted as `Cluster CR KDE`, to the corresponding radial color distribution around reference random points denoted as `Field CR KDE`, and to the radial galaxy profile around refernce random points in the wide field survey $P(R \,|\, \lambda, z)$ denoted as `Field R KDE`. From the deep field feature dataset two KDEs are constructed, one covering all features with their full value ranges $P(m, \vec{c}, \vec{s}, z_g)$ denoted as `DEEP Full KDE`, and one containing only the distribution of wide field colors restricted to the wide field depth $P(\vec{c}_{ref}$ denoted as `DEEP C KDE`. For the full deep field KDE the magnitude range is set as $i < 25.5$, which is done to avoid edge-effects in estimating the KDE, however in the future when drawing random samples we always restrict the drawing range to $i < 24.5$.

In the above, folowing section 4.3.3.2, we only considered a set of reference colors $\vec{c}_{ref}$ for the wide field, which are a subset of the available colors $\vec{c}$. These were selected to bracket the approximate location of the D4000 break at the targeted cluster redshift range, and are listed in Table 4.4.

4.3.4.2 Survey Resampling

In the following step we perform the statistical background subtraction and survey extrapolation via resampling, as defined in section 4.3.3.3.

First, a proposal distribution is defined, from the combination of the `DEEP Full KDE` and `FIELD R KDE` modules, which is given by Equation 4.25, and denoted as `Full Proposal` in Figure 4.4.

From the full proposal distribution a random sample is drawn

$$\vec{\theta}_i = (m_i, \vec{c}_i, \vec{s}_i, z_{g;i} R_i)\,, \tag{4.30}$$

and a corresponding restricted random vector is defined using a subset of the values of the respective elements

$$\vec{\theta}_{i;\mathrm{wide}} = (\vec{c}_{i;\mathrm{wide}}, R_i)\,. \tag{4.31}$$

The probability of this restricted draw is then evaluated with three different distributions: The `Cluster Target Score` and the `Field Target Score` are found by evaluating $\vec{\theta}_{i;ref}$ with `Cluster CR KDE` and `Field CR KDE` respectively. The `Restricted Proposal Score` is found by evaluating $\vec{\theta}_{i;ref}$ with a restricted proposal distribution defined based on the

modules DEEP C KDE and FIELD R KDE (both of which is restricted to $i < 22.5$) following Equation 4.26. This probability is then multiplied by the factor $\mathcal{M} = 100$, as defined in section 4.3.3.3.

From the above three scores two ratios are taken: The ratio of the Cluster Target Score and Proposal Score corresponds to the right hand side of Equation 4.28 and is denoted as Cluster Ratio (CL), and the ratio of Field Target Score and Proposal Score correspond to the left hand side of Equation 4.27 and is denoted as Field Ratio (FL).

Combining the above with a uniform random draw u_i drawn from $\mathcal{U}[0; 1)$ indicated as Uniform Draw allows us to evaluate the inequality given by Equation 4.27 and Equation 4.28 for the random draw $\vec{\theta}_{i;ref}$ and the corresponding $\vec{\theta}_i$. This inequality corresponds to the decision denoted as the element FL < U < CL in Figure 4.4. For random draws where this criterion is fulfilled, $\vec{\theta}_i$ is accepted as a sample from the extrapolated cluster member distribution and collated to the dataset indicated as Cluster Members. For random draws which do not fulfill FL < U < CL, a further decision is taken corresponding to the inequality defined in Equation 4.29, indicated as the element U < FL. Random draws $\vec{\theta}_i$ which fulfill this criterion are accepted as samples from the extrapolation of the field distribution, and are collated in the dataset indicated as Field Objects. Random draws which fulfill neither of these criteria are discarded, and a new random draw is performed to get $\vec{\theta}_{i+1}$.

Using the above approach, the random draws can be repeated until a sufficiently large sample is accepted for the cluster member and the field object dataset. These accepted draws can either be used directly to construct mock observations, or alternatively a KDE can then be constructed to estimate the p.d.f. of the cluster members and extrapolated field galaxies, denoted as Cluster Member KDE and Extrapolate Field KDE. Estimating the extrapolated feature distribution of cluster member galaxies was the target of this work, and is achieved via the above algorithm. A further outcome of this implementation is that a cluster member KDE is constructed, no further p.d.f. evaluations are necessary, and subsequent samples can be drawn directly from the final cluster member KDE. Although this cluster member KDE incurs additional smoothing compared to its input samples, it enables a drastic reduction in computational cost in generating further samples, enabling the efficient scalability of the algorithm.

In addition, the consistency of the extrapolated field distribution estimate with the wide field and deep field galaxy distributions provides a useful null-test for judging the performance of the method.

4.3.4.3 Nested Spherical Cluster Model

The above sections provide the recipe to estimate the Cluster Member KDE. However, a practical limitation of the method is that since the proposal R_i values are drawn from the full considered radial range around clusters and reference random points (see Table 4.3) the larger radial ranges will be much better sampled than the lower radius ranges because of the increase in surface areas $p(R) \propto R$. Due to this geometric behaviour, ensuring that the cluster member distribution is well sampled at all radii becomes computationally expensive.

A simple way around this problem is to split the radial range into a set of radial bins, and

radial bin	radial range [arcmin]	padding range [arcmin]
1	$[10^{-1.5};\ 10^{-0.5})$	$[10^{-1.5};\ 10^{0})$
2	$[10^{-0.5};\ 10^{0})$	$[10^{-1.5};\ 10^{0.5})$
3	$[10^{0};\ 10^{0.5})$	$[10^{-1.5};\ 10^{1})$
4	$[10^{0.5};\ 10^{1})$	$[10^{-1.5};\ 10^{1.2})$

Table 4.5: Nested Sphere definitions for the radial bins

repeat the resampling for each bin separately. To avoid edge-effects in the KDEs caused by the sharp radial bin edges, instead of disjoint radial ranges we define a set of four nested concentric spheres. Each sphere covers a target radial range, but extends from a lower limit out to a padding radius beyond the upper limit of the target radius, as listed in Table 4.5. However, when drawing samples, only those within the target radial range are considered for the resampling. Therefore while the resampling is performed separately for the different radial ranges, consistency between the KDEs is enforced by the fact that they are derived from shared information due to the radial overlap.

This approach repeats the resampling scheme four times, yielding estimates for both the `Cluster Member KDE` and the `Extrapolated Field KDE` for the different radial ranges. While each of these p.d.f.-s is individually normalized to unity, we express the relative probability p_l of a member galaxy residing in a given radial bin r_l around a cluster as

$$p_l \approx \frac{\hat{n}_{c;\,l} - \hat{n}_{r;\,l}}{p_l(i < 22.5)} \Big/ \sum_l \frac{\hat{n}_{c;\,l} - \hat{n}_{r;\,l}}{p_l(i < 22.5)} \tag{4.32}$$

where $\hat{n}_{c;\,l}$, $\hat{n}_{r;\,l}$ is the average number of galaxies around clusters and random points residing in the radial bin in the wide field dataset, and $p_l(i < 22.5)$ is the probability that based on the `Cluster Member KDE` in radial bin l a galaxy is bright enough to be in the wide field selection. While this formalism is similar to the direct background subtraction scheme defined is section 4.3.3.1, it is only used to approximate the relative weight of different radial ranges, and does not influence the estimation of the feature p.d.f.-s within the radial ranges.

4.4 Model Results

In this section we present our statistical model for the different sets of galaxy populations measured from the DES data, and overview the properties of the estimated cluster member galaxy population which is obtained as the result of the statistical background subtraction and survey depth extrapolation. These task are performed with the aim of creating an *emulator* for the galaxy distribution, that is to construct a statistical model which enables generating random realizations of the underlying galaxy populations. This approach enables us to subsequently generate mock observations.

In addition to the above, as a sanity check we also overview the properties of the extrapolated galaxy feature distribution measured around random points, to test the consistency of the statistical resampling method defined in section 4.3.

4.4.1 Input Feature Distributions

The primary inputs for the measurement presented in this paper are the distribution of galaxy properties around galaxy clusters and reference random points in the DES wide field survey (section 4.2.1), and the distribution of galaxy properties in the DES deep fields (section 4.2.3).

Throughout the statistical modelling process presented in section 4.3, galaxy properties are collated into *features*, which are listed in Table 4.1. The dependence of wide field galaxy features on the distance from clusters and reference random points is measured during the analysis, and from the deep field and wide field feature dataset a feature probability distribution function is estimated using kernel density estimates. For each targeted sample of galaxy clusters a total of five different p.d.f.-s are estimated as listed in Table 4.3. Of these distributions in the following we present the distribution measurements and the corresponding KDE estimates for the two primary distributions: The distribution of features around clusters in the wide field data (`Cluster CR KDE`), and the distribution of all measured features in the deep field dataset (`Deep Full KDE`). This restriction is motivated as the distribution estimates for `Field CR KDE` and `Field R KDE` measured in the same framework as `Cluster CR KDE` but around reference random points, while `Deep C KDE` is comprised of a subset of `Deep Full KDE`.

4.4.1.1 Distributions of wide field galaxies around clusters

Figure 4.7 shows the measured feature distribution of galaxies around a selection of redMaPPer galaxy clusters with $\lambda \in [45; 60)$ and $z \in [0.3; 0.35)$. The features of this distribution are the reference colors $\vec{c}_{ref} = (g - r, r - i)$ and the projected radial separation R measured from the target galaxy cluster centers. Using these sets of features a KDE is constructed according to section 4.3.2, whose model for the p.d.f is shown as the continuous curves and contours on Figure 4.7. Visualizing the full joint distribution of these three features is non-trivial, therefore we only show a selected set of slices (conditional distributions), and the corresponding KDE model. We note that the KDE is constructed globally for all features and the full value range, and not only for the shown conditional distributions. The top left two panels of Figure 4.7 show the distributions of galaxy colors at different projected radii from the cluster centers for all galaxies with $i < 22.5$. The histograms correspond to the measured distribution, while the continuous curve represents the appropriate slice of the global KDE model. It is visible that there is a prominent radial dependence as the red sequence becomes increasingly dominant for small radii. This represents the increased prevalence of red cluster member galaxies in the line-of-sight, as blue cloud galaxies become a subdominant component. The KDE model provides a good overall description of these galaxy distributions capturing the two-component nature of the galaxy population. It recovers the position and the approximate relative weight of the red sequence population, and also approximates the extreme peak of the red sequence distribution. We note that since the targeted galaxy clusters span a redshift range $\Delta z = 0.05$, the width of the

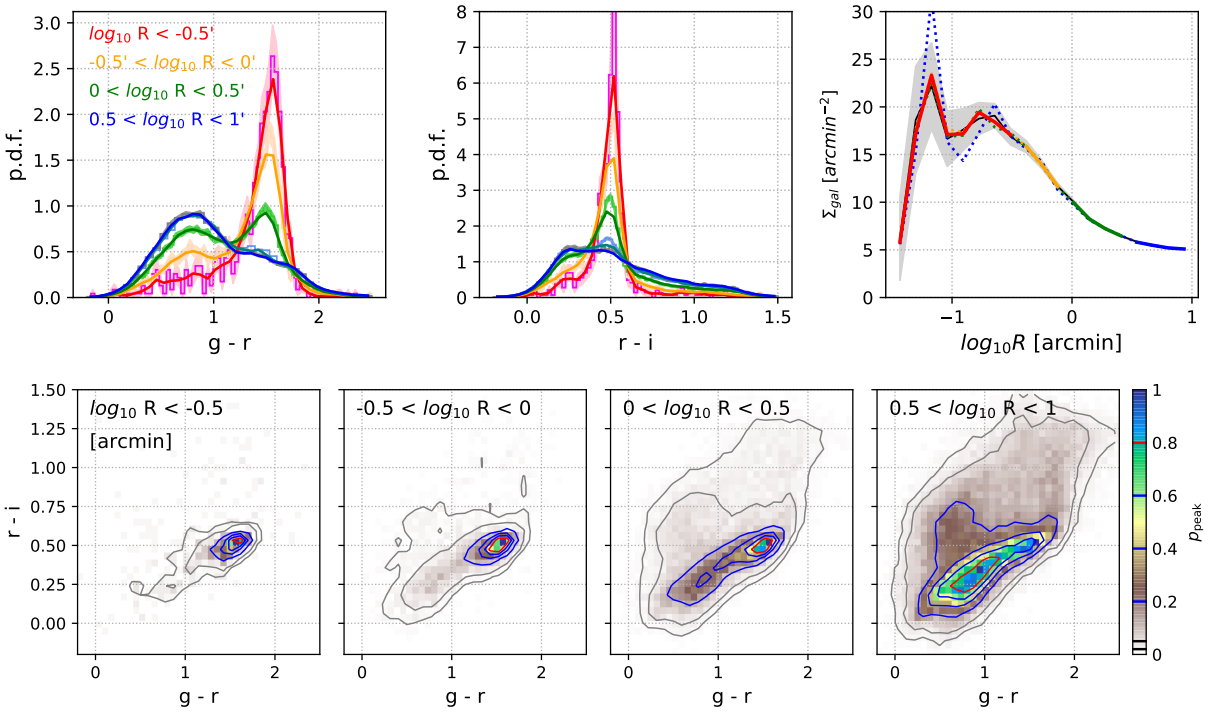

Figure 4.7: Distribution of galaxy features with $i < 22.5$ around redMaPPer galaxy clusters ($\lambda \in [45; 60)$, $z \in [0.3; 0.35)$ in the DES wide field dataset . **Top left and center:** $g - r$ and $r - i$ color histograms of galaxies in bins of projected radius. *histogram:* DES data. *contours:* KDE reconstruction. The radial bins correspond to the radial shells defined in Table 4.5 **Top right:** Surface density profile of galaxies around the targeted cluster sample. *black:* measured profile. *color:* KDE reconstruction of the surface density profile, color coded to the radial bins of the the top left and center panels. **Bottom:** $g - r$ - $r - i$ color distribution of galaxies in the four radial shells. *histogram:* DES data. *contours:* KDE reconstruction. We note that the KDE is constructed globally for the full magnitude and feature ranges, and not only for the shown 2d marginal distribution.

observed red sequence population is measured to be wider, by this dispersion, compared to its intrinsic width.

The top right panel of Figure 4.7 shows the surface density profile of galaxies with $i < 22.5$ around the selected cluster sample in the wide field survey as the solid black curve, and colored curves show the corresponding KDE models for the four nested shells according to section 4.3.4. In addition to the target range of the KDEs which are shown as the full colors, as a consistency test the interior continuation of the KDE model for each nested spherical bin (Table 4.5) is shown as the dotted lines, of which only the one corresponding to the outermost radial range shows a mild deviation from the respective profile of the data. It is visible that for the wide field galaxy selection used in this study (see Table 4.2) the cluster represents a very significant overdensity, with the surface density profile declining with radius and reaching the survey average at large radii.

The measured radial surface density profile and the KDE models show very good agreement. This is a strong indication of the performance of the KDE representation of the survey data, since the KDE method estimates the absolute number density of galaxies, and not the the surface number density. Recovering these two profiles is a markedly different task, as due to the increase in surface area element with radius, the number density profile possesses an asymptotic power-law density profile, while the surface density profile remains within the same order of magnitude at all radii. Therefore the excellent agreement between the measured and recovered surface density profile means that the deviation between the measured and modeled density is small over a density range of several orders of magnitude (as set by the change in area element).

The bottom panels of Figure 4.7 show the $g - r$ - $r - i$ color-color diagram of galaxies with $i < 22.5$ in the same four radial bins as used in the upper panels. The histograms correspond to the measured distributions, while the contours represents the appropriate slice of the global KDE model and correspond to 2%, 5%, 20%, 40%, 60% and 80% of the peak of the conditional probability density. It is visible that using a global fit for all three features and the full parameter ranges, the KDE estimate is able to at the same time recover both the central behaviour of the distributions (blue and red contours compared with green and blue histogram cells), and the wings of the distributions (gray contours compared with gray and brown histogram cells).

4.4.1.2 Distributions of deep field galaxies

Figure 4.8 shows the $g - r$ - $r - i$ and the $r - i$ - $i - z$ color - color diagrams of the deep field galaxies in four different magnitude ranges. The measured distribution is shown as a histogram, and the corresponding KDE model is represented by the contours defined as 2%, 5%, 20%, 40%, 60% and 80% of the peak of the conditional p.d.f. shown. In a similar manner, Figure 4.9 shows the color-magnitude diagrams measured with the above colors and i-band magnitudes. It is visible that for both figures, the global KDE model is able to simultaneously provide an excellent description for both the center and also for the wings of the distribution.

Figure 4.10 shows the distribution of features related to the morphology parameters of the `ngmix` galaxy model. The left panel shows the marginal distribution of the size proxy feature $\log_{10}(1 + T)$ with i-band magnitude, where T is the effective area of the galaxy measured in arcseconds (see Table 4.1). We note that the appearance of negative T values is related to noise incurred during the `ngmix` photometry fit, as very small and zero size objects (such as stars) sometimes are fitted with nominally negative sizes. Naturally, when rendering such objects the render-size is set to 0. The measured distribution is shown as a histogram, and the corresponding KDE model is represented by the contours defined as 2%, 5%, 20%, 40%, 60% and 80% of the peak of the conditional p.d.f. shown. In a similar manner, the right panel of Figure 4.10 shows the corresponding distribution between estimated absolute galaxy ellipticity $|g|$ and i-band magnitude. For both cases the measured distributions and the KDE model of the p.d.f-s show excellent agreement.

The center panel of Figure 4.10 shows the histograms of the bulge/disk ratio of the `ngmix` galaxy model for two magnitude bins $19.5 < i < 21$ and $21 < i < 22.5$, along with the corresponding KDE model. The panel shows that brighter galaxies are more likely to be bulge dominated (e.g. described by a De Vaucouleur light profile) compared to fainter galaxies, which is in ac-

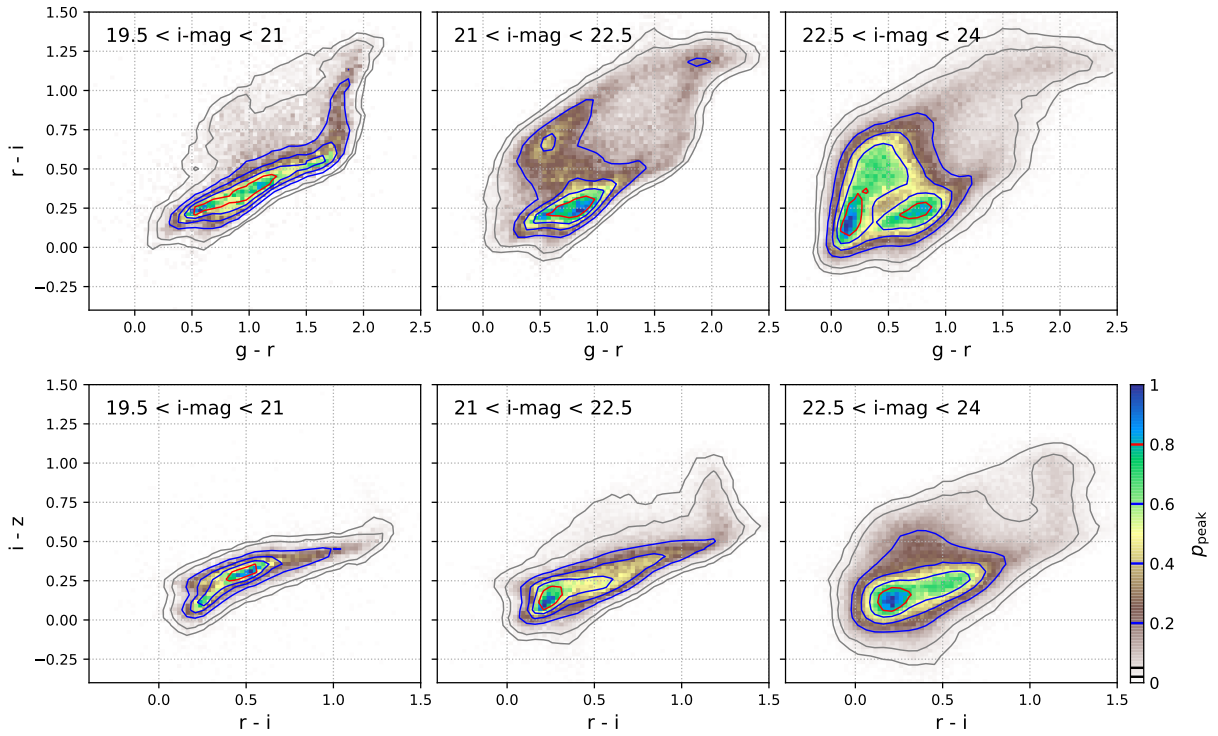

Figure 4.8: Distribution of $g-r$, $r-i$, $i-z$ galaxy colors in the DES deep field dataset in bins of i-band magnitude. *histogram:* DES data. *contours:* KDE reconstruction. We note that the KDE is constructed globally for the full magnitude and feature ranges, and not only for the shown 2d marginal distribution.

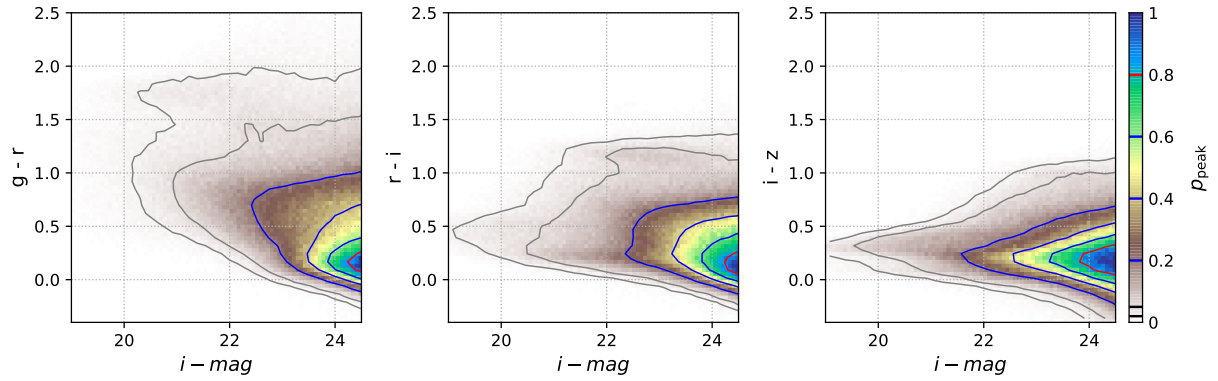

Figure 4.9: Distribution of the $g-r$, $r-i$, $i-z$ galaxy colors against i-band magnitude magnitudes in the DES deep field dataset. *histogram:* DES data. *contours:* KDE reconstruction. We note that the KDE is constructed globally for the full magnitude and feature ranges, and not only for the shown 2d marginal distribution.

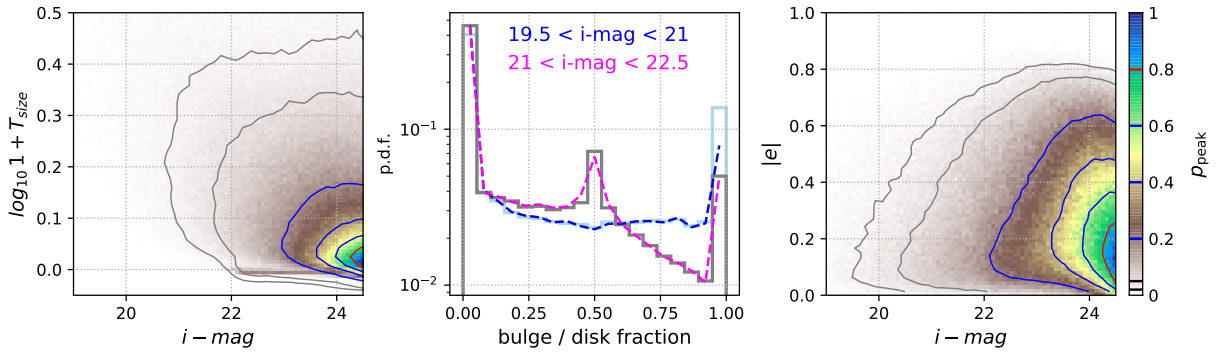

Figure 4.10: Distribution of galaxy morphology parameters in the DES deep field dataset, as listed in Table 4.1. *histogram:* DES data. *contours / curves:* KDE reconstruction. We note that the KDE is constructed globally for the full magnitude and feature ranges, and not only for the shown marginal distributions.

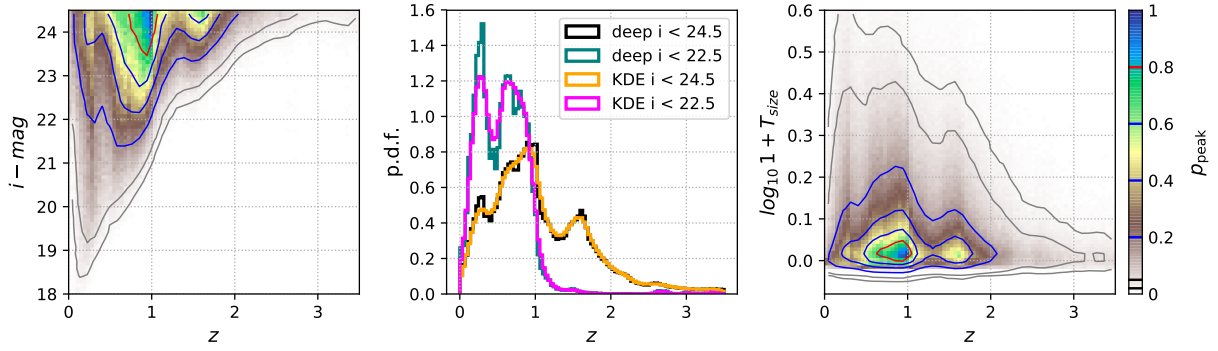

Figure 4.11: Redshift distributions of deep field galaxies estimated by the EAzY algorithm for the DES dataset (section 4.2.3), along with the correlation of *i*-band magnitude and size proxy against redshift. *histogram:* DES data. *contours / curves:* KDE reconstruction. We note that the KDE is constructed globally for the full magnitude and feature ranges, and not only for the shown marginal distributions

cordance with expectations from galaxy evolution (Gavazzi et al., 2010). The peak appearing at 0.5 is an imprint of the morphology prior of the deep field photometry pipeline, and it becomes prominent for the fainter galaxy selection as there the available information to constrain morphology from survey observations diminishes. KDE estimates cannot reproduce the hard cutoff edges [0; 1] of the bulge/disk value ratios, and for this reason we cap the distributions around 0 and 1 to restrict the p.d.f. model to the appropriate interval, so that values greater than 1 or lower than 0 receive a value of 1 or 0 respectively. Nevertheless, with this modification, the KDE model is able to provide a good description of the observed bulge / disk distributions.

Figure 4.11 shows the estimated redshift distribution of the deep field galaxies, as predicted by the EAzY algorithm (Brammer et al. 2008 see section 4.2.3). The center panel shows the marginal redshift p.d.f. along with the KDE reconstruction for two different magnitude ranges, and the left and right panels show the correlation of photometric redshift estimate with *i*-band magnitude and galaxy size respectively. As expected, the photometric redshift algorithm predicts that fainter galaxies are at higher redshifts, and that higher redshift galaxies appear on average smaller.

4.4.2 Cluster Member Feature Distributions

We now discuss the output of the statistical modeling presented in section 4.3, which is the primary result of this study. In short, the model performs a statistical background subtraction to isolate the cluster member galaxies of redMaPPer selected galaxy clusters. It extrapolates the cluster member population to fainter magnitudes by reconstructing the cluster member galaxy distribution from deep field galaxies, thereby keeping the additional features measured on better quality data in the deep field catalog. The end products of these operations is a set of random samples drawn from the feature p.d.f. of the cluster member galaxies, and a set of random samples which are drawn from the extrapolated field galaxy population. For both of these samples a KDE is constructed according to section 4.3.2, whose purpose is to provide a computationally efficient way of generating samples (albeit at the cost of increased smoothing), since the full rejection sampling algorithm carries a large computational cost due to the multiple KDE scorings and low acceptance rate. The resulting models allow us to estimate in a statistical way the full structure of a line-of-sight, along with the population of cluster member galaxies, foreground and background galaxies, and accounting for the redshift distributions of the various components.

The above statistical modeling is performed for a set of six cluster selections arranged in richness as $\lambda \in [30; 45)$ and $\lambda \in [45; 60)$ and redshift $z \in [0.3; 0.35)$, $z \in [0.45; 0.5)$ and $z \in [0.6; 0.65)$ as shown on Figure 4.2, and it is the models for these samples which we describe in detail below.

4.4.2.1 Line-of-Sight Model

A key outcome of the statistical background subtraction and survey extrapolation scheme is a spatial and photometric model for an observed galaxy cluster line-of-sight. Using the emulated model for the feature distribution and relative abundance of cluster member galaxies, and the emulated model feature distribution of the foreground and background galaxies (field galaxies)

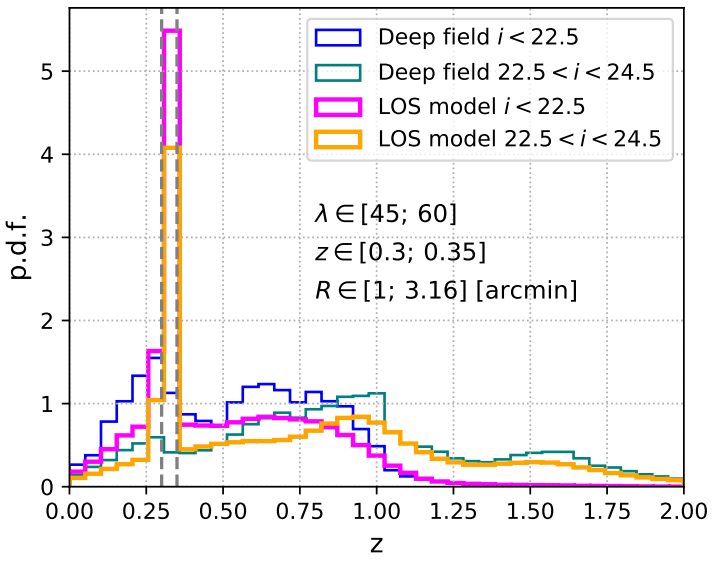

Figure 4.12: Line of sight model for the redshift distribution of galaxies near clusters with $\lambda \in [45; 60)$ and $z \in [0.3; 0.35)$ within the projected radial range $R \in [1; 3.16)$. *(magenta, orange)* redshift distribution model around clusters in different magnitude bins. *(blue, green)* photometric redshift distribution measured in the DES deep fields in different magnitude bins. *grey dashed* limits of the cluster redshift range. It is visible that the cluster contains a significant fraction of galaxies in the line of sight.

we can construct the line of sight as a combination of these two terms. In this line of sight model the cluster member galaxies are located at the mean cluster redshift, while the line of sight galaxies inherit the redshift p.d.f. specific for their deep field observed features. This line-of-sight galaxy distribution model is illustrated by the redshift distributions shown on Figure 4.12 for a cluster sample with $\lambda \in [45; 60)$ and $z \in [0.3; 0.35)$ where the emulated redshift p.d.f of galaxies with $i < 22.5$ and within the radial range $R \in [1; 3.16)$ arcmin is shown as the magenta histogram, which is a combination of a cluster member term located at $z = 0.325$, and a field term. As comparison the redshift p.d.f. of deep field galaxies is shown in blue for the same magnitude range. Owing to the extrapolation part of the analysis, the reconstructed line of sight is modeled down to the deep field limiting magnitude of $i < 24.5$. It contains a faint cluster member population in addition to the faint end of the field galaxy population shown as the orange histogram, with the comparison redshift distribution of the deep field galaxies is shown as the green histogram.

The above line-of-sight model therefore incorporates the appropriate redshifts at which galaxies reside, and allows us to correctly take the lens geometry into account when constructing a mock observation and apply an appropriate redshift dependent lensing effect onto the galaxies. It is furthermore visible in Figure 4.12 that the redshift distribution of galaxies near a cluster in projection is significantly different from the one in the deep fields, due to the prominent overdensity at the cluster redshift. This aspect of the line-of-sight model enables us to construct mock

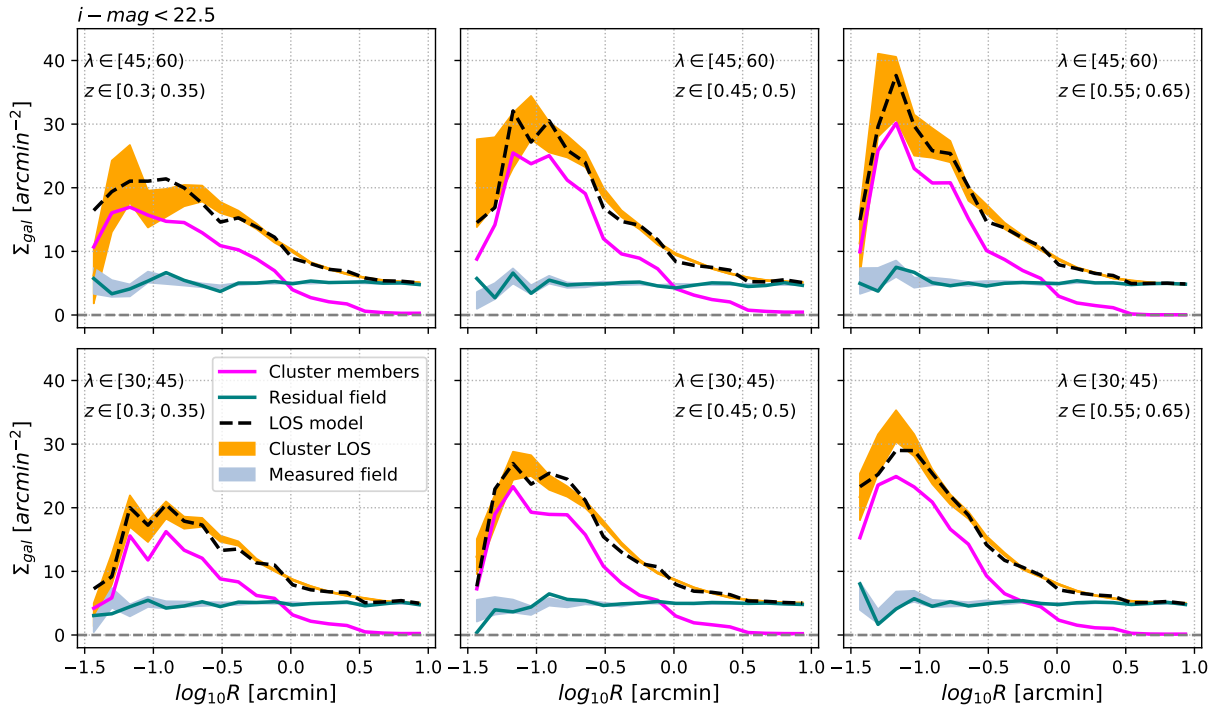

Figure 4.13: Line of sight model for the surface density of galaxies around galaxy clusters with different richness and redshift. *(orange)* Surface density profile measured around redMaPPer clusters. The width of the shaded area represents the Poisson uncertainty propagated into surface density. *(gray)* Surface density of galaxies measured around reference random points. *(green)* model for the surface density profile of *field* galaxies within the cluster line-of-sight. *(magenta)* model for the surface density profile of *cluster member* galaxies in the cluster line-of-sight. *(black dashed)* model for the total galaxy surface density profile in the cluster line-of-sight (the sum of the green and magenta curves).

observations where we can test the impact of the galaxy cluster on the performance of photometric redshift estimates. In cluster weak lensing measurements this manifest itself as the problem of *boost factors* or cluster member contamination as introduced in chapter 3, as well as propagatig blending-related photometry effects onto the performance estimates of photometric redshifts.

Figure 4.13 shows the measured and modeled galaxy surface density profiles for the six cluster richness and redshift samples considered in this study. To enable comparison with the surface density profiles measured directly in the DES wide field survey the magnitude range is restricted to $i < 22.5$. For each cluster sample (each panel) we show the measured projected galaxy surface density around the clusters in DES, indicated by the orange lines, and the surface density profile around the corresponding sample of reference random points as the gray lines. The width of these lines indicates the Poisson uncertainty of the number of galaxies in each radial bin, propagated into an error on the surface density. It is visible that at large radii the measured surface density profile around clusters approaches the approximately uniform surface density distribution measured around the reference random points.

The model for the field population is shown as the green lines on Figure 4.13. This distribution corresponds to the background model during the statistical background subtraction, but it is constructed from the re-weighting and resampling of deep field galaxies in section 4.3.4. The excellent agreement between this and the profile measured around random points in the DES wide field data is a strong consistency test of the the statistical model, and is an indication that the statistical background subtraction is performed correctly.

The model for the pure cluster member distribution is shown as the magenta curves on Figure 4.13. It is visible that this component captures the radial variations in surface density, and this density profile approaches zero at large radii, consistent with the finite extent of the cluster galaxy populations. The model for the full surface density profile is then obtained as the sum of the cluster member (magenta) and the field (green) population estimates, and this surface density profile is shown as the black dashed lines, which can then be directly compared with the galaxy profiles measured in the DES data around clusters (orange lines). It is visible that overall the two lines agree well. The downturn of the surface density profiles at low radii is explained by the drop in galaxy detection efficiency due to the higher galaxy density. Obscuration and blending is expected to play an increased role, and the presence of the bright cluster galaxy (BCG) and the intra-cluster light (ICL) is expected to reduce the performance of photometry.

4.4.2.2 Galaxy Distributions for Clusters and Field

The distribution of photometric features such as colors and magnitudes of the different galaxy populations and their relative weights are also modeled during the background subtraction and feature extrapolation, as presented in section 4.3. Galaxy clusters host a characteristic population and overdensity of quiescent red galaxies distributed along the red-sequence. These galaxies, along with the non-red cluster member component and the foreground and background galaxies, constitute the galaxies in a line-of-sight. As presented in section 4.4.2.1, galaxy clusters manifest as an overdensity compared to the median galaxy surface density on the sky. This overdensity and the corresponding radial profile is different for the different galaxy populations, color and magnitude ranges. Throughout this analysis we measure and model the projected feature distribution around clusters, as presented in section 4.4.1.1, and the essence of the applied method is to decouple this into a model on the feature distribution (in this case of color and magnitude) of *cluster member* galaxies, and foreground/background (i.e. *field*) galaxies.

Figure 4.14, Figure 4.15 and Figure 4.16 represent the measurements and the models corresponding to the color and magnitude distribution of galaxies considered in this analysis for the three higher richness cluster samples $\lambda \in [45; 60)$ $z \in [0.3; 0.35)$, [0.45; 0.5) and [0.6; 0.65) respectively. Given that the relevant subspace of the feature space spans four dimensions: two colors, i-band magnitude and projected radius, we can only show a subset of selected conditional distributions in print. For the three cluster redshift ranges we show galaxy distribution in $g-r$, $g-r$ and $r-i$ colors respectively. There are three magnitude ranges shown for each cluster sample (panel), the first two [19; 21) and [21; 22.5) is fitted to the DES wide field data, while the third [23; 24) is a pure extrapolation based on the algorithm.

The color distributions measured in the corresponding magnitude and radial range around galaxy clusters in the DES wide field data are shown as the orange histograms, with the height

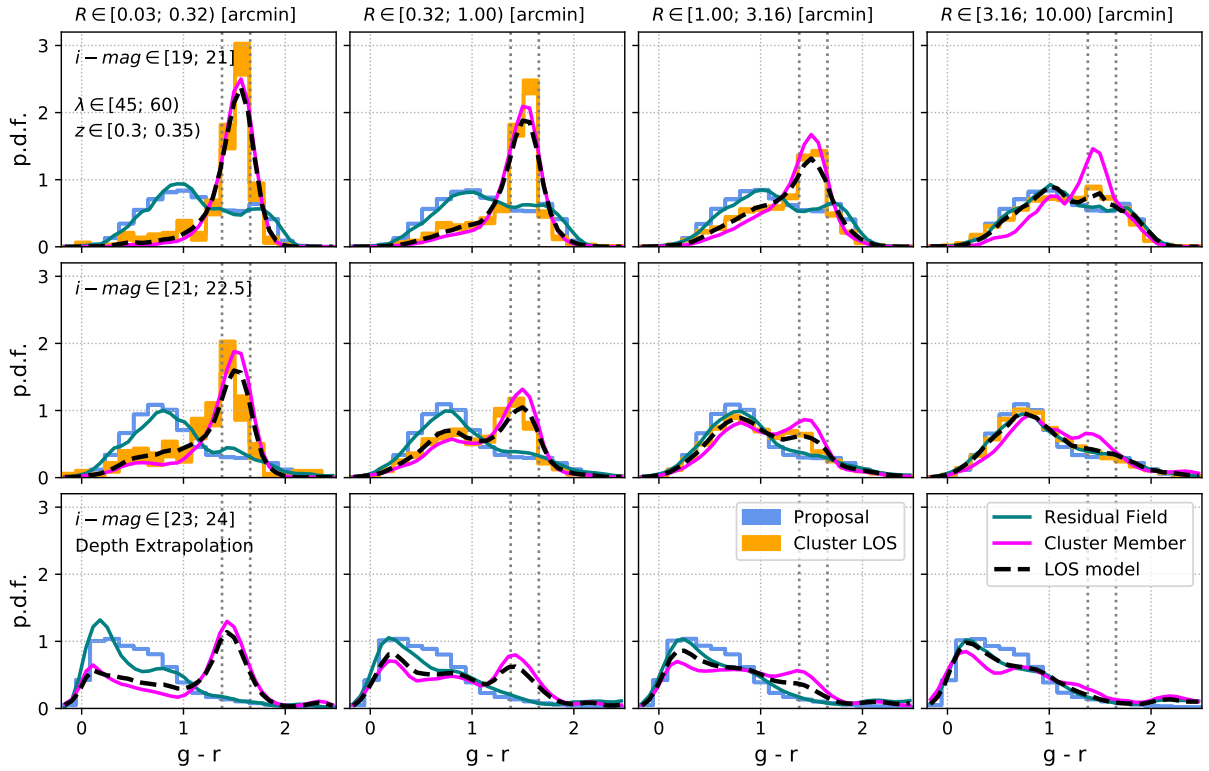

Figure 4.14: Conditional color distribution of galaxies around galaxy clusters across four projected radial regimes (shown in the different columns) around galaxy clusters with $\lambda \in [45; 60)$ and $z \in [0.3; 0.35)$. The distribution of galaxies are shown in $g-r$, $g-r$ and $r-i$ colors respectively. There are three magnitude ranges shown (rows), the first two [19; 21) and [21; 22.5) is fitted to the DES wide field data, while the third [23; 24) is a pure extrapolation based on the algorithm. **Orange**: color p.d.f. measured as a histogram around galaxy clusters in DES data. The height of the shaded area indicates the Poisson uncertainty propagated into the normalized histogram. **Blue**: color distribution measured within the corresponding magnitude range in the DES deep fields. This distribution is identical for each column and for all cluster samples. **Green**: Model for the color distribution of foreground and background galaxies in the line of sight. **Magenta**: Model for the color distribution of cluster member galaxies. **Black dashed**: Model for the full line of-sight, which can be directly compared with the orange histogram. **Gray dotted**: 1σ location of the redMaPPer red-sequence cluster member galaxies.

of the colored area representing the Poisson uncertainty of the measurement. As a comparison for each cell the respective conditional color distribution measured in the DES deep fields is shown (blue histogram). This population naturally has no radial dependence, and therefore the histogram is identical in the different columns of each panel. Due to the large number of deep field galaxies, the noise in the deep-field histogram is negligible. Out of the above two populations, only the deep field one is measured down to the third magnitude bin $i \in [23; 24)$, therefore the orange cluster measurement is not shown there. It is visible that the color distribution around clusters shows a strong radial trend, with the orange histogram approaching the blue with increasing radius. A dominant driver of this trend is increasing prominence of the red-sequence at low radii, which manifests as a peak in the color distribution. It is also visible that the relative weight of the red-sequence is greater for brighter galaxies, and the difference between cluster and field line of sights is also greater for brighter galaxies. As a reference, the location of the redMaPPer red-sequence model is indicated by the vertical gray dotted lines on Figure 4.14, Figure 4.15 and Figure 4.16. These lines correspond to the 1σ range of the membership probability weighted color distribution of redMaPPer cluster members for that cluster richness, redshift and galaxy magnitude range. It is visible that both the location and also the width of the peak of the orange histogram coincides with the the location of the redMaPPercluster members, indicating that it is indeed an imprint of the red-sequence. We note that in the current case the width of the red-sequence is determined not only by its intrinsic width, but also by the redshift range of the cluster sample which broadens it compared to that at a single redshift. The indicated width of the red-sequence is calculated considering the full magnitude range. The location of the peak of the measured cluster line-of-sight shows a consistent slight shift between the brighter and fainter magnitude range, which is indicative of the slight slope of the red-sequence.

Figure 4.14, Figure 4.15 and Figure 4.16 show the emulator *model* for the galaxy distributions around galaxy clusters as the black dashed lines, which can be directly compared with the orange histogram of the corresponding raw measurement. This model is derived without direct information about the wide field galaxy luminosity function around clusters, and only using information from the deep field data. Nevertheless, as visible on the upper two rows of Figure 4.14, Figure 4.15 and Figure 4.16, the line of sight model can describe the magnitude dependent color variations of the galaxy distributions, and well approximate the relative weight of the red-sequence peak, albeit while slightly over-estimating its width. The bottom row of each panel shows the model for galaxies in the line of sight with $i \in [23; 24)$. Due to the extrapolation part of the approach, the model extends for these fainter magnitudes, even though the cluster lines-of-sight are not measured. For all three redshift ranges, the extrapolated model differs from the deep field color distribution, by increasing the probability of the redder colors. However only for the highest ($z \in [0.6; 0.65)$) and lowest ($z \in [0.3; 0.35)$) redshift range does this take the form of a well defined peak at the continuation of the red-sequence. This shortcoming is likely related to the bandpass coverage of the 4000 angstrom break at the $z \in [0.45; 0.5)$ redshift range by $g-r$ and $r-i$ colors, and given the good performance in the two other redshift ranges is likely not an intrinsic limitation of the extrapolation method.

The above described model for the full line-of-sight is expressed as a non-negative linear

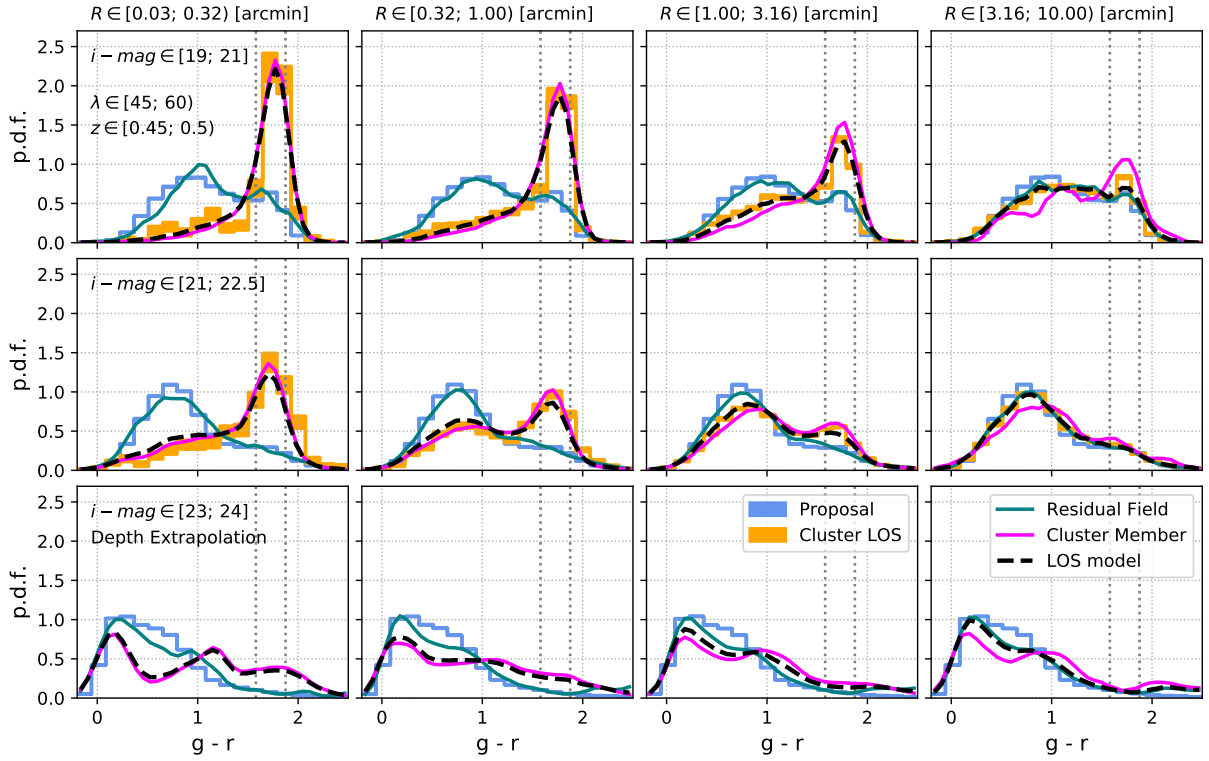

Figure 4.15: Conditional color distribution of galaxies around galaxy clusters across four projected radial regimes (shown in the different columns) around galaxy clusters with $\lambda \in [45; 60)$ and $z \in [0.45; 0.5)$.The distribution of galaxies are shown in g–r, g–r and r–i colors respectively. There are three magnitude ranges shown (rows), the first two [19; 21) and [21; 22.5) is fitted to the DES wide field data, while the third [23; 24) is a pure extrapolation based on the algorithm. **Orange**: color p.d.f. measured as a histogram around galaxy clusters in DES data. The height of the shaded area indicates the Poisson uncertainty propagated into the normalized histogram. **Blue**: color distribution measured within the corresponding magnitude range in the DES deep fields. This distribution is identical for each column and for all cluster samples. **Green**: Model for the color distribution of foreground and background galaxies in the line of sight. **Magenta**: Model for the color distribution of cluster member galaxies. **Black dashed**: Model for the full line of-sight, which can be directly compared with the orange histogram. **Gray dotted**: 1σ location of the redMaPPer red-sequence cluster member galaxies.

combination of a *cluster member* and a *residual field* component,[4] indicated by the magenta and the green curves respectively in Figure 4.14, Figure 4.15 and Figure 4.16. Both of these components are derived according to the resampling approach presented in section 4.3.3.3, and thus they are not directly estimated from data. These two components correspond to the similarly color coded lines on Figure 4.13, and while the cluster member galaxies are placed at the mean cluster redshift of the bin, the galaxies in the residual field population are distributed along the full line-of-sight.

The distribution and properties of foreground and background galaxies is independent of the cluster galaxy population. For this reason it is expected for the model for the residual field to be independent of radius. Furthermore, while the bright tip of the DES deep fields is not fully representative of the actual median DES wide field survey line of sight due to sample variance, they still provide a reasonable reference distribution. Comparing the residual field model (green curve) with the deep field distribution (blue histogram) on Figure 4.14, Figure 4.15 and Figure 4.16, it is visible that there are no strong radial variations in the distributions, and that the residual field indeed approximates the deep field distribution, with only minor deviations visible at the faint end of the panels (lowest row on each panel).

Finally, the color distribution model for cluster member galaxies (magenta curve on Figure 4.14, Figure 4.15 and Figure 4.16) shows both a significant radial trend, and a strong deviation from the residual field component across all cluster richness and redshift samples and radial ranges. It is visible that the cluster members contain a prominent red-sequence component, whose strength decreases with radius and magnitude. Due to the overdensity of the cluster, at low radii cluster member galaxies dominate the line of sight, and their distribution closely resembles the color distribution of all line-of-sight galaxies (as most galaxies are actually in the cluster). With increasing radius the cluster member distribution becomes bluer, yet a significant fraction of red galaxies remain in the brighter magnitude range.

The radial color evolution of the cluster member galaxy population can be described by the approximate red-fraction, whose radial profile for the six cluster richness and redshift bins is shown in Figure 4.17. The regions of the radial color-color space used defining galaxies as "red" and "blue" are indicated by the black dashed lines in Figure 4.18 for the three different cluster redshift samples. As visible these regions are chosen to bracket the position of the red sequence which is dominant at low radii. and this way estimate the approximate red-fraction of the galaxy sample as function of radius.

Figure 4.17 shows the approximate red-fraction for two magnitude ranges, a brighter bin covering $i \in [19; 22.5)$ shown in magenta, which coincides with the DES wide field depth, and a fainter bin covering $i \in [22.4; 24.5)$ shown in orange which is derived from a purely extrapolated color-color distribution. Profiles from the higher richness cluster sample are shown as the solid lines, while the profiles from the lower richness samples are shown as the dashed lines. There appears to be no significant difference between the richness bins. The bright galaxy sample shows a clear monotonic trend in all redshift and richness samples, where the red fraction decreases from approximately unity at very low projected radii to approximately 30% - 40% at

[4]We denote the *residual field* as such, as it is the field population estimate obtained from the resampling and extrapolation, and it is not directly fitted to the deep field dataset.

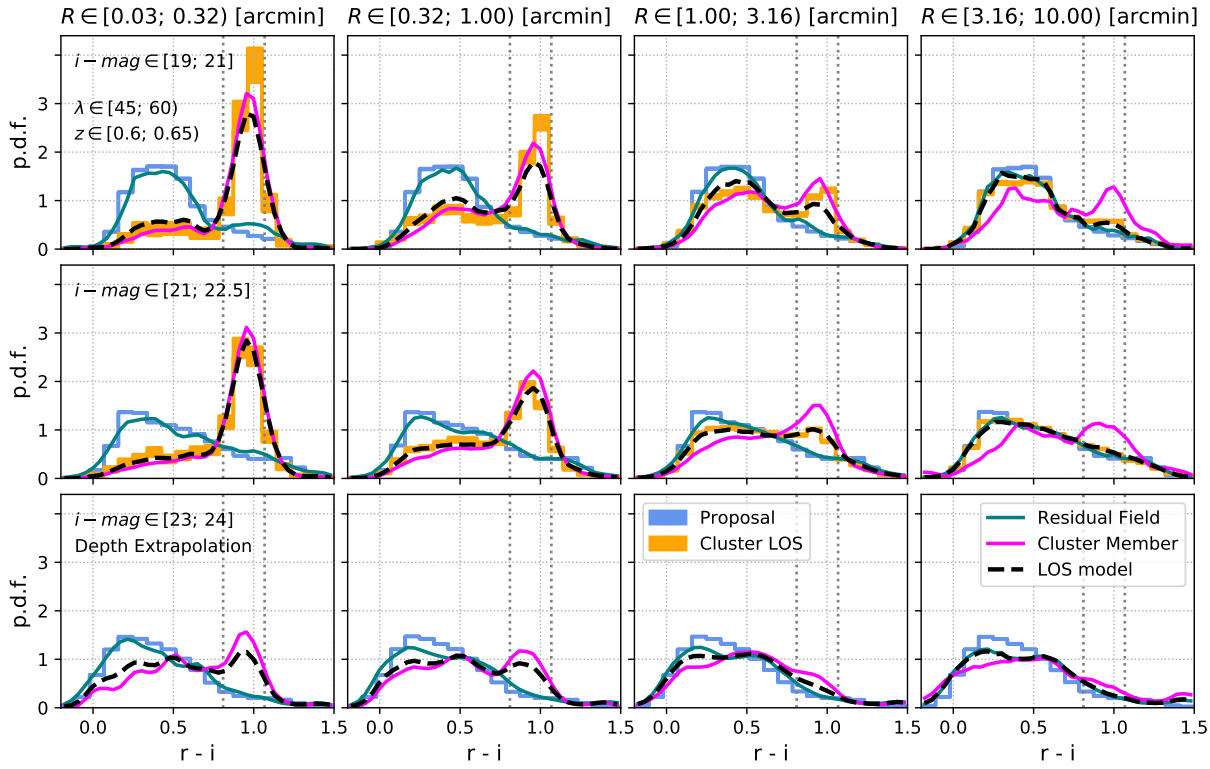

Figure 4.16: Conditional color distribution of galaxies around galaxy clusters across four projected radial regimes (shown in the different columns) around galaxy clusters with $\lambda \in [45; 60)$ and $z \in [0.6; 0.65)$. The distribution of galaxies are shown in $g-r$, $g-r$ and $r-i$ colors respectively. There are three magnitude ranges shown (rows), the first two [19; 21) and [21; 22.5) is fitted to the DES wide field data, while the third [23; 24) is a pure extrapolation based on the algorithm. **Orange**: color p.d.f. measured as a histogram around galaxy clusters in DES data. The height of the shaded area indicates the Poisson uncertainty propagated into the normalized histogram. **Blue**: color distribution measured within the corresponding magnitude range in the DES deep fields. This distribution is identical for each column and for all cluster samples. **Green**: Model for the color distribution of foreground and background galaxies in the line of sight. **Magenta**: Model for the color distribution of cluster member galaxies. **Black dashed**: Model for the full line of-sight, which can be directly compared with the orange histogram. **Gray dotted**: 1σ location of the redMaPPer red-sequence cluster member galaxies.

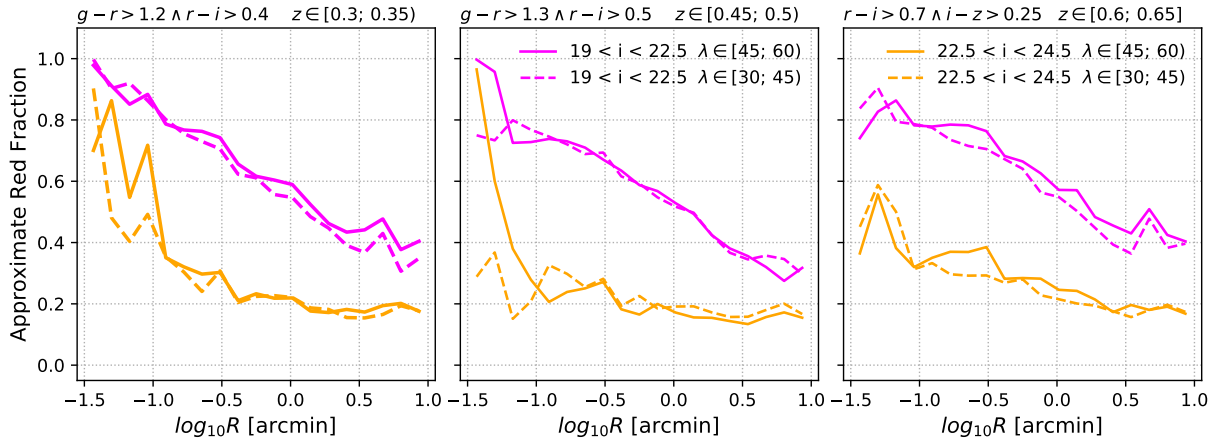

Figure 4.17: Approximate red fraction of cluster members as a function of projected radius for six different cluster richness and redshift samples. *(magenta)* estimated red fraction for the wide-field magnitude range. *(orange)* estimated red fraction derived from the depth extrapolated cluster member distribution. The definition of the color cells are listed at the top of each panel.

large radii approaching 10 arcmin. This qualitative behaviour is consistent with theoretical expectations (Butcher & Oemler, 1978; Hansen et al., 2009) and the measurement from numerical simulations presented in chapter 3 (see Figure 3.2). The same behaviour is not uniformly true for the fainter red-fraction profiles, where while some cluster samples show a prominent red galaxy population at the center, the decline is much faster for these fainter populations than the brighter counterparts for the same clusters, and at large radii the galaxy population appears to show lesser color gradients, and appear to be bluer.

4.4.2.3 Photometric Redshifts of Cluster Members

In chapter 3 of this thesis we highlighted a systematic effect originating from mis-estimating the photometric redshifts for cluster member galaxies. Resulting from that, cluster members can enter the background source galaxy catalog and bias the lensing measurement. As discussed there, this is primarily a result of using the redshift p.d.f of galaxies in the median cosmic line of sight as representative for galaxies with the additional selection in the vicinity of a galaxy cluster at cluster redshift $z_{\rm clust}$.

Using the presented line-of-sight model for cluster members and field galaxies we can illustrate the full extent of this type of bias. The redshift information in this model originates from the deep field data and is estimated by the `EAzY` algorithm (Brammer et al., 2008). In the cluster and line-of-sight model this photometric redshift is used only for the foreground and background galaxies, while cluster members are placed at the cluster redshift. Nevertheless it is illuminating to inspect what redshifts are predicted for cluster members. One can formulate this as the question: Given the connection between features and redshift in the deep fields, what would we conclude on the redshift distribution of cluster member galaxies? This form of prior or training based inquiry is the core of the photometric redshift algorithms used in current surveys (Bonnett

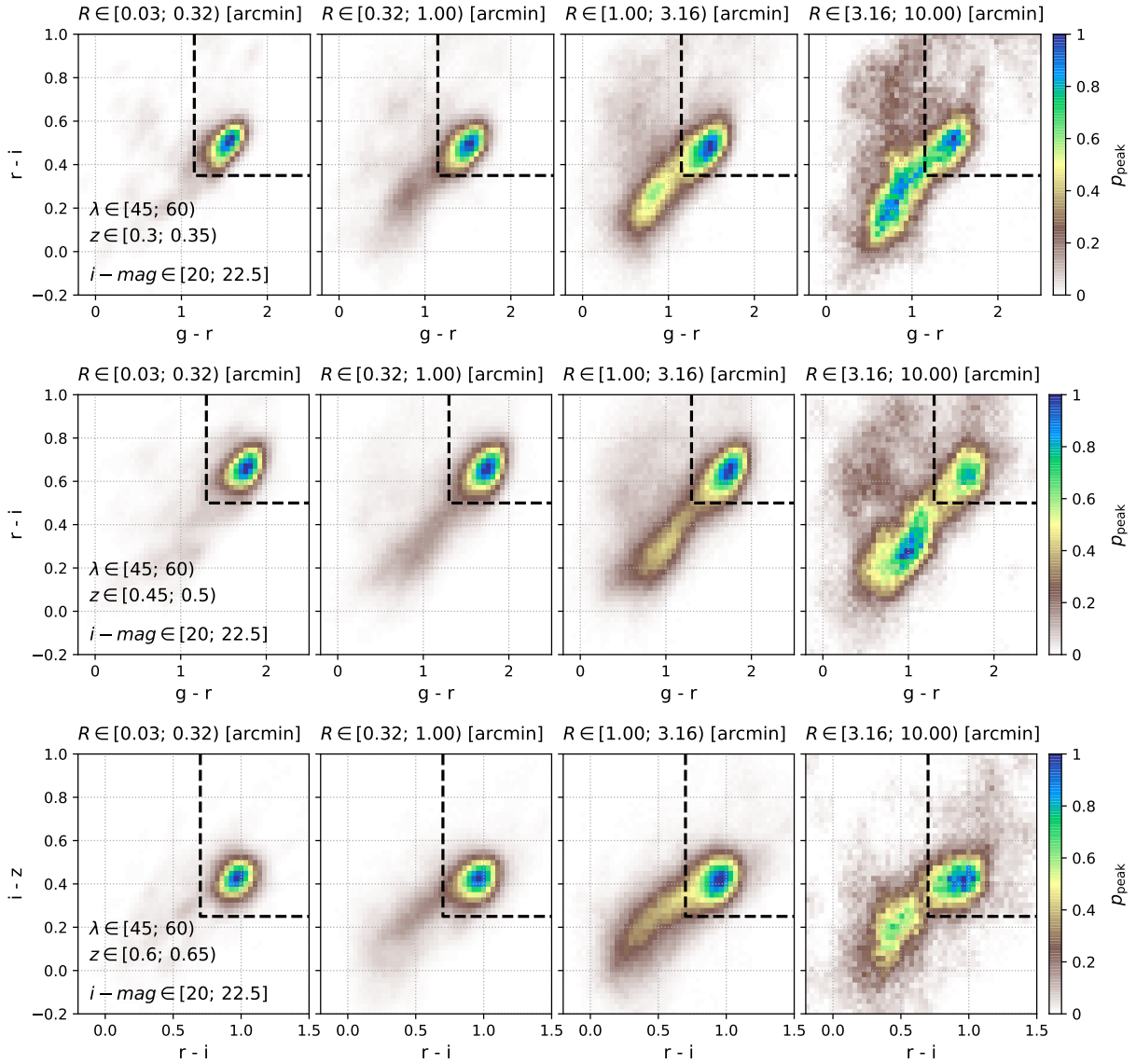

Figure 4.18: Color-color diagrams of the cluster member components for the three higher richness cluster samples located at different redshifts. The region delimited by the black dashed lines indicate the areas used to calculate the approximate red fraction in Figure 4.17

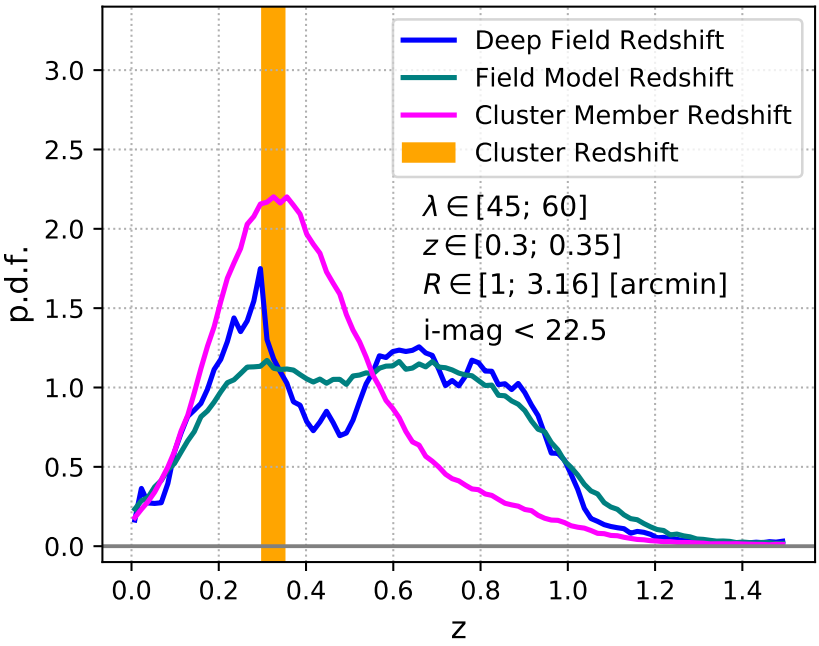

Figure 4.19: Photometric redshift distributions extrapolated from the deep field dataset for the cluster member and foreground and background galaxies in the line-of-sight model.

et al., 2016; Hoyle et al., 2018). As visible, it does not take information about the fact that we are considering objects with a specific selection function, namely cluster member galaxies.

The photometric redshift distributions predicted for the various line-of-sight components are illustrated on Figure 4.19 for galaxies around clusters with $\lambda \in [45;\ 60]$ and $z \in [0.3;\ 0.35]$. There the blue line shows the photometric redshift p.d.f. of deep field galaxies, while the green line shows the modeled redshift p.d.f. for the foreground and background galaxies in the line-of-sight. This latter p.d.f follows the deep field p.d.f. well, which is expected to an extent since the deep field observations are chosen and constructed to be a test dataset for the typical galaxy properties and distributions in the wide field data. The location of the targeted galaxy clusters in redshift is indicated by the vertical orange region, and the photometric redshift p.d.f. assigned to the cluster member galaxies is shown as the magenta curve. It is visible that the cluster member redshift p.d.f. is markedly different from both the deep field and also from the foreground and background p.d.f. and shows a prominent peak centered at the redshift range of the targeted clusters with a tail extending to higher redshifts.

In reality, all member galaxies are located within $z \in [0.3;\ 0.35]$, and the difference between that range and the magenta curve is representative of the uncertainty of photometric redshifts of cluster member galaxies. The above shown behaviour drives the leakage of cluster members into the background source sample, and ends up biasing the cluster mass measurements. While the study in chapter 3 used a specific redshift related selection function of the weak lensing analysis, the above results can be approximately compared with the redshift p.d.f. shown on Figure 3.5, which shows similar widths and shape for the redshift p.d.f. of the cluster galaxy component.

Technically, for cluster member galaxies a more precise photometric redshift could be derived by fitting a red sequence model. This is however adding extra information about the SED of the galaxies. The above noisy redshift estimate instead assumes that the distribution of galaxy SEDs and redshifts in the deep field is representative of the targeted galaxy population.

4.5 Generating Random Lines-of-Sight

The previous sections described the results of the statistical model for the galaxy content of galaxy clusters, and the line of-sight distributions of the foreground and background galaxies. The aim of this section is then to generate mock cluster images where each is representative of the galaxy content of the mean cluster within a given range of cluster observables, and also contain the mean line-of-sight within which clusters reside in the survey. To achieve this, first a draw of non-central galaxies is performed (section 4.5.1), then the gravitational shear and magnification effects are applied in section 4.5.2, then the model for the central galaxy and the intra-cluster light is constructed section 4.5.3, and finally the resulting cluster model is rendered into a survey-like exposure using a simplified version of the DES galsim image simulation pipeline in section 4.5.4.

To render galaxies into images we make use of the `ngmix` bulge + disk light profile model (Sheldon, 2015) whose parameters are the position, magnitudes (fluxes), effective size T, shape $(g_1; g_2)$ and the bulge / disk fraction of the galaxy. These parameters are derived from the features of the statistical model used in this analysis (see Table 4.1).

4.5.1 Random Draws of Galaxy Populations

In the statistical model (section 4.3) the distribution of non-central galaxies are represented by probability distribution functions, which includes two main components: the distribution of cluster member galaxies (satellites) and the distribution of foreground and background galaxies in the respective lines-of-sight. Using these, a synthetic cluster and line-of-sight can be created through random draws from the p.d.f. of the different components. Here each draw corresponds to adding a new galaxy to a mock field-of-view catalog with an angular and redshift position, and the photometric and morphological features contained within the model. In the following we describe the details on constructing mock cluster fields-of-view at the level of these mock catalogs.

The p.d.f. representation of the galaxy distributions do not carry information about the absolute number of objects, therefore this needs to be set based on the observed number of galaxies. In real observations only the bright end of the luminosity function is observed in the survey (i.e. $i < 22.5$) therefore the number of fainter galaxies must be defined according to their relative probability in the model p.d.f. Furthermore, the galaxy distribution model is divided into four wide bins in projected radius as described in section 4.3.4.3 and listed in Table 4.5 to improve the numerical performance of the calculations.

Based on the above, constructing a mock galaxy catalog is performed the following way:

1. For each radial range l, calculate $\hat{N}_{C;l}$ and $\hat{N}_{R;l}$ the mean number of galaxies with $i < 22.5$ around clusters and random points respectively in radial range l listed in Table 4.5.

2. Estimate the number of objects belonging to the cluster, and to foreground and background objects based on Equation 4.32. Furthermore, to discretize the expected galaxy number, for each radial range we take a Poisson random number of galaxies based on the mean number as

$$N_{M;l} = \text{Poisson}\left(\frac{\hat{N}_{C;l} - \hat{N}_{R;l}}{p_{\text{memb};l}\,(i < 22.5)}\right), \tag{4.33}$$

and

$$N_{R;l} = \text{Poisson}\left(\frac{\hat{N}_{R;l}}{p_{\text{rand};l}\,(i < 22.5)}\right). \tag{4.34}$$

3. Draw cluster members $N_{M;l}$ times from $p_{\text{memb};l}$ and foreground and background galaxies $N_{R;l}$ times from $p_{\text{rand};l}$.

4. For cluster members set the redshift as z_{clust}.

5. Convert the projected radius feature R_i into 2D position assuming circular symmetry in a flat-sky approximation via

$$\text{RA}_i = R_i\ \cos\phi_i \tag{4.35}$$
$$\text{DEC}_i = R_i\ \sin\phi_i\,, \tag{4.36}$$

where ϕ_i is drawn from $\mathcal{U}[0;\ 2\pi)$, and where the mock cluster is located at $\text{RA}_{\text{target}} = 0$ and $\text{DEC}_{\text{target}} = 0$.

The outcome of the above recipe is a galaxy catalog which contains cluster members and for- and background galaxies each distributed according to their respective statistical models derived from the survey data, but extrapolated to a fainter limiting magnitude, and the surface density of galaxies is set to the mean surface density measured around galaxy clusters.

A useful regularization of the above method is to update step 1 by only measuring $\hat{N}_{C;l}$ from data, and expressing $\hat{N}_{R;l}$ as a function $\hat{N}_{C;l}$ using the statistical model. In practice this is achieved by taking the ratio of accepted events during the rejection sampling (see section 4.3.3.3) which only fulfill Equation 4.29, to the amount of events which fulfill both Equation 4.29 and Equation 4.28. This latter formulation avoids scenarios when due to measurement noise by chance $\hat{N}_{R;l} > \hat{N}_{C;l}$.

Lastly, we note that the above approach considered constructing galaxy clusters as a sum of independent draws of non-central galaxies. In case of a complicated cluster finder and selection function, this assumption might be incorrect as the cluster finder might rely on additional correlations and connections within the cluster member galaxies.[5] Hence the above approach is admittedly only the simplest way to construct mock clusters, and if required the resulting mock galaxy distributions can be further filtered based on the selection function of a cluster finder.

[5]Technically independent random draws could produce a mock cluster, albeit at low likelihood, where every galaxy is much fainter than the mean brightness of the parent distribution. Such an outlier mock cluster might not be detected by cluster finders

4.5.2 Cluster Lens Model and Galaxy Shapes

The essence of these simulations is to benchmark weak lensing measurements, and this requires us to define a mass model for the galaxy cluster and apply the appropriate gravitational shear to background galaxies. For this we make use of the mass models and mass constraints found in chapter 2 of this thesis. The present analysis uses a richness binning scheme which is a subset of that in chapter 2. Furthermore, as that analysis did not find a significant redshift evolution in the richness-mass scaling, we can approximate the relevant mean cluster masses for the present mocks with the appropriate values listed in Table 2.2, that is $M_{200\mathrm{m}} \approx 10^{14.45}\ M_\odot$ for the $\lambda \in [30; 45)$ bin and $M_{200\mathrm{m}} \approx 10^{14.65}\ M_\odot$ for the $\lambda \in [45; 60)$ bin across the three different redshift bins.

In the following we only consider the mass model for the 1-halo term which is dominant on the small scales considered in this study, and consists of a spherically symmetric mass distribution with Navarro-Frenk-White (NFW) mass profile (Navarro et al., 1996, also see Equation 2.33). This lens mass distribution is placed at the cluster redshift z_{clust} and subsequently gravitational shear and magnification is applied to line-of-sight galaxies based on their true redshifts assigned by the model. The lensing effect induced by a NFW halo can be expressed analytically following (Oaxaca Wright & Brainerd, 1999), or by inverting the approach followed in section 2.5.1, to obtain the projected surface mass density profile $\Sigma(\theta)$ and the excess surface mass density profile $\Delta\Sigma(\theta)$ (defined according to section 1.3).

From these density profiles we express the convergence experienced by the image of the i-th source galaxy with redshift z_i and angular separation θ_i from the cluster center as

$$\kappa_i = \begin{cases} \Sigma(\theta_i) \,/\, \Sigma_{\mathrm{crit}}(z_{\mathrm{clust}},\, z_i) & z_i > z_{\mathrm{clust}} \\ 0 & z_i \leq z_{\mathrm{clust}} \end{cases} \tag{4.37}$$

where $\Sigma_{\mathrm{crit}}(z_{\mathrm{clust}},\, z_i)$ is the critical surface density defined in Equation 1.21. In a similar manner, the tangential shear is given by

$$\gamma_{\mathrm{T};\,i} = \begin{cases} \Delta\Sigma(\theta_i) \,/\, \Sigma_{\mathrm{crit}}(z_{\mathrm{clust}},\, z_i) & z_i > z_{\mathrm{clust}} \\ 0 & z_i \leq z_{\mathrm{clust}}\,. \end{cases} \tag{4.38}$$

In practice, the quantities of interest are the reduced tangential shear $g_{\mathrm{T};\,i} \equiv \gamma_i \cdot (1 - \kappa_i)^{-1}$ and the magnification $\mu_i \equiv \left((1-\kappa_i)^2 - \gamma^2_{\mathrm{T};\,i}\right)^{-1}$. A shape parameter of this form $g_i = (g_{1;\,i};\ g_{2;\,i})$ is indeed an input in the `ngmix` galaxy model, therefore this formulation can be directly used in creating mock galaxy light profiles. The magnification is however only applied as the simplest possible approximation, by modulating the total flux of the galaxy light models $F_{\mathrm{lensed};\,i} = \mu_i\, F_i$ in an a-chromatic way. This is technically incorrect, as magnification does not directly increase the total flux, but increases the apparent sizes of sources at fixed pre-seeing surface brightness.

In the above we described the approach gravitational lensing impacts the shape and flux of galaxies, but these quantities must be simulated together with the intrinsic shape of galaxies $\mathrm{e}_{\mathrm{gal}}$. In this model we assume no preferential intrinsic alignment and therefore the shape of each

$z \in$	$\lambda \in$	$\langle i \rangle$	$\langle g-r \rangle$	$\langle r-i \rangle$	$\langle i-z \rangle$	$\langle T \rangle$ [arcsec^2]	$\langle \lvert g \rvert \rangle$
[0.3; 0.35)	[30; 45)	17.76	1.36	0.54	0.32	28.90	0.14
[0.3; 0.35)	[45; 60)	17.62	1.38	0.54	0.31	33.20	0.14
[0.45; 0.5)	[30; 45)	18.58	1.85	0.70	0.37	21.92	0.15
[0.45; 0.5)	[45; 60)	18.50	1.85	0.71	0.37	28.43	0.14
[0.6; 0.65)	[30; 45)	19.36	1.83	1.01	0.44	16.90	0.17
[0.6; 0.65)	[35; 60)	19.18	1.83	1.02	0.45	22.44	0.16

Table 4.6: Properties of the mean bright central galaxy (BCG) across the different cluster richness and redshift bins. For each BCG the bulge (DeVaucouleours) fraction is set to unity.

galaxy is determined by the absolute ellipticity $\lvert e \rvert_i$ as defined by the feature $\vec{s}_1$ in Table 4.1. From this the shape of each galaxy is given by

$$\mathbf{e}_{\mathrm{gal};\,i} = \begin{pmatrix} e_{\mathrm{gal};\,1} \\ e_{\mathrm{gal};\,2} \end{pmatrix}_i = \begin{pmatrix} -\lvert e \rvert_i \cdot \cos(2\,\phi_i) \\ -\lvert e \rvert_i \cdot \sin(2\,\phi_i) \end{pmatrix} \tag{4.39}$$

where ϕ_i is a random angle drawn from $\mathcal{U}[0,\ \pi)$.

From the combination of the gravitational reduced shear and the intrinsic shape of the object we can calculate a net shape distortion Jacobian for each galaxy, which transforms the intrinsically circular light model into the desired shape and alignment as

$$A_{\mathrm{net};\,i} = A_{\mathrm{grav};\,i} \cdot A_{\mathrm{intr};\,i} = (1-\kappa) \begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix} \cdot \begin{pmatrix} 1-e_{gal;\,1} & -e_{gal;\,2} \\ -e_{gal;\,2} & 1+e_{gal;\,1} \end{pmatrix} \tag{4.40}$$

where the first two terms refer to the weak lensing Jacobian expressed for reduced shear, while the rightmost term relates to the intrinsic ellipticity and alignment of the galaxy. From the above matrix product the net e_1 and e_2 values can be derived as

$$e_{\mathrm{net};\,1} = g_1 + e_{\mathrm{gal};\,1} - g_1 e_{\mathrm{gal};\,1} - g_2 e_{\mathrm{gal};\,2} \approx g_1 + e_{\mathrm{gal};\,1} \tag{4.41}$$

$$e_{\mathrm{net};\,2} = g_2 + e_{\mathrm{gal};\,2} - g_1 e_{\mathrm{gal};\,2} + g_2 e_{\mathrm{gal};\,1} \approx g_2 + e_{\mathrm{gal};\,2}\,, \tag{4.42}$$

and these in turn can then be directly used as an input for the intrinsically circular `ngmix` galaxy light distribution model used in rendering the galaxies into images.

4.5.3 BCG and Intra-Cluster Light Model

A prominent feature of galaxy clusters is the presence of a *bright central galaxy* (BCG) and a surrounding distribution of *intra-cluster light* (ICL) emitted by a diffuse stellar component bound in the cluster halo. These components contain a significant fraction of the total optical light emitted by the cluster, therefore accounting for them is essential in a dedicated simulation of synthetic galaxy cluster observations (Zhang et al., 2019a; Kluge et al., 2020).

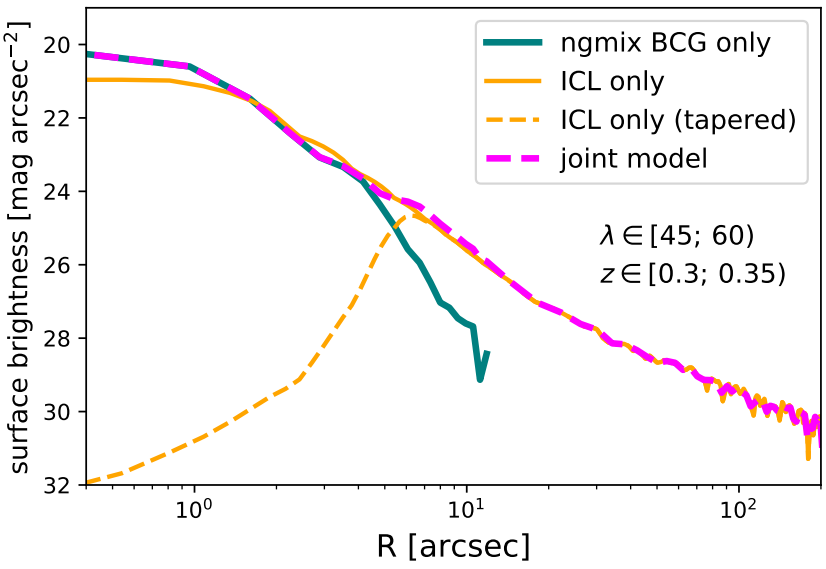

Figure 4.20: Surface brightness model for the intra-cluster light (ICL) and bright cluster galaxy (BCG). The joint model is found as the sum of the BCG and a tapered ICL model.

By construction galaxy clusters identified by redMaPPer are always centered on a bright red-sequence galaxy. This is a simplified view of reality, as in recent mergers or in non-equilibrium systems the more cental galaxy might not be red or the brightest, or there might be multiple similarly bright BCGs (Rykoff et al., 2014). Nevertheless originating from the special location they inhabit, the central galaxies of massive halos follow a different evolutionary track compared to satellite galaxies. It is observed that their properties are closely tied to the mass and properties of their cluster (Postman & Lauer, 1995), and their luminosity function is approximately Gaussian at fixed cluster properties (Hansen et al., 2009). Based on these observations we model the synthetic central galaxy in the mocks as having the mean properties of the redMaPPer central galaxies in the cluster sample. The relevant mean central galaxy features are listed in Table 4.6 for the different cluster redshift and richness samples. The central galaxies are assumed to have a De Vaucouleours light profile, and the only stochastic element in the model is their random orientation in the plane of the sky with fixed ellipticity $|g|$.

The total light in the central region of a cluster is however not fully described by the above model, as there is a continuous transition between the light usually associated with the central galaxy and the intra-cluster light (Kluge et al., 2020). Zhang et al. (2019a) investigated the properties of the ICL for redMaPPer selected galaxy clusters in DES with $z_{\mathrm{clust}} \in [0.2; 0.3)$, by conducting a stacked analysis mapping the the diffuse light of the ICL down to surface brightness of 30 mag arcsec^{-2}. Furthermore, they investigated the richness (mass) dependence of the ICL, finding a self-similarity of the light profile when expressed in units of R_{200m}. Using these findings Gruen et al. (2019) constructed a simple model for the ICL observed around redMaPPer clusters in DES. This model extrapolates from the measurement of Zhang et al. (2019a) in terms of cluster mass using the self-similarity of the profiles, and also in terms of cluster red-

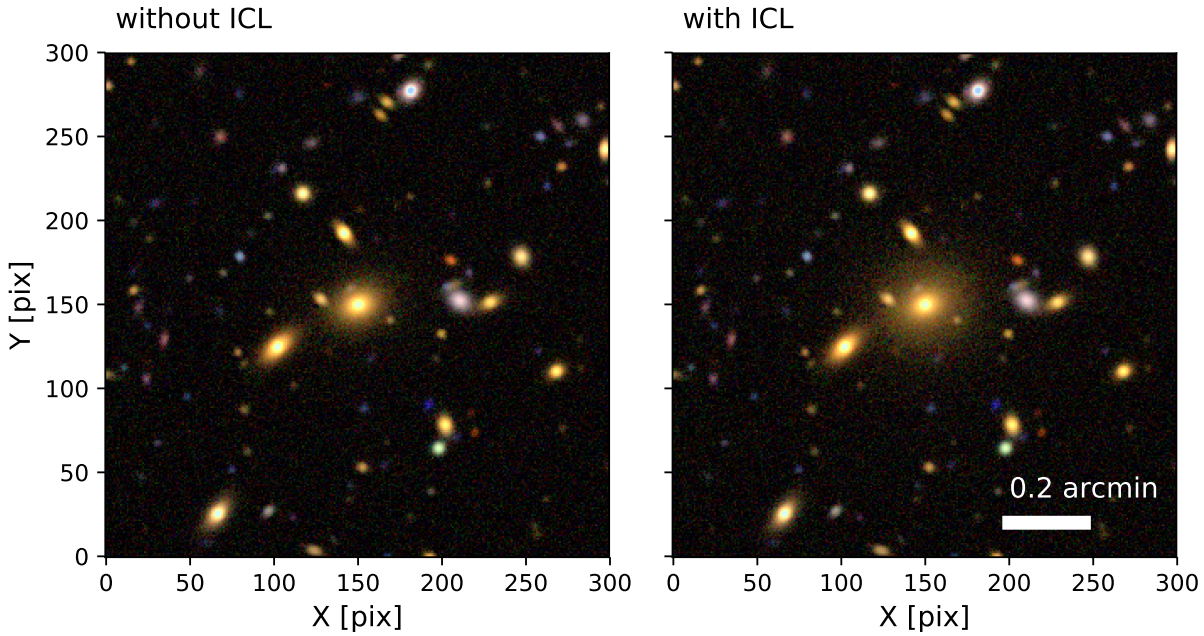

Figure 4.21: Synthetic center of a mock galaxy cluster without (left) and with the intra-cluster light model applied (rigth), as defined in Figure 4.20

shift by assuming a simple passively evolving stellar population within the ICL. We note that this latter assumption is closely related to the formation history and age of the, ICL which is poorly constrained from current observational studies due to the difficulty of high redshift observations. Hence in case of a late-forming ICL the above extrapolation overestimates the total light contained in it at early times. Furthermore the model neglects the mild radius dependent color gradient in the ICL, where the outer ranges are slightly bluer.

In the following we adopt the above ICL model of Gruen et al. (2019). As a simplification we assume that the colors of the ICL are identical to the mean colors of BCGs at that redshift and cluster richness sample. The remaining task is to construct a joint light model for the BCG model and smooth ICL component which can be efficiently rendered into mock images. The light profiles of these two components are illustrated on Figure 4.20: The ICL component extends to large radii as an approximate power law surface density light profile, while the `ngmix` BCG light model is dominant in the inner regions. Because of their overlap, these components cannot be directly added to each other. Therefore we define a tapered ICL model where the tapering scale is set by the size of the BCG component $\theta_S = \sqrt{T_{BCG}}$. Using a simple empirical sigmoid tapering we define a joint light profile model

$$\mu(\theta) = \mu_{BCG}(\theta) + \left(1 - \frac{1}{1 + e^{2\,(\theta - \theta_S)}}\right) \mu_{ICL}(\theta)\,, \tag{4.43}$$

which is shown on Figure 4.20 as the magenta dashed curve. As visible this joint model captures both the small scale and the large scale behaviour of the light profile, while smoothly connecting the two components.

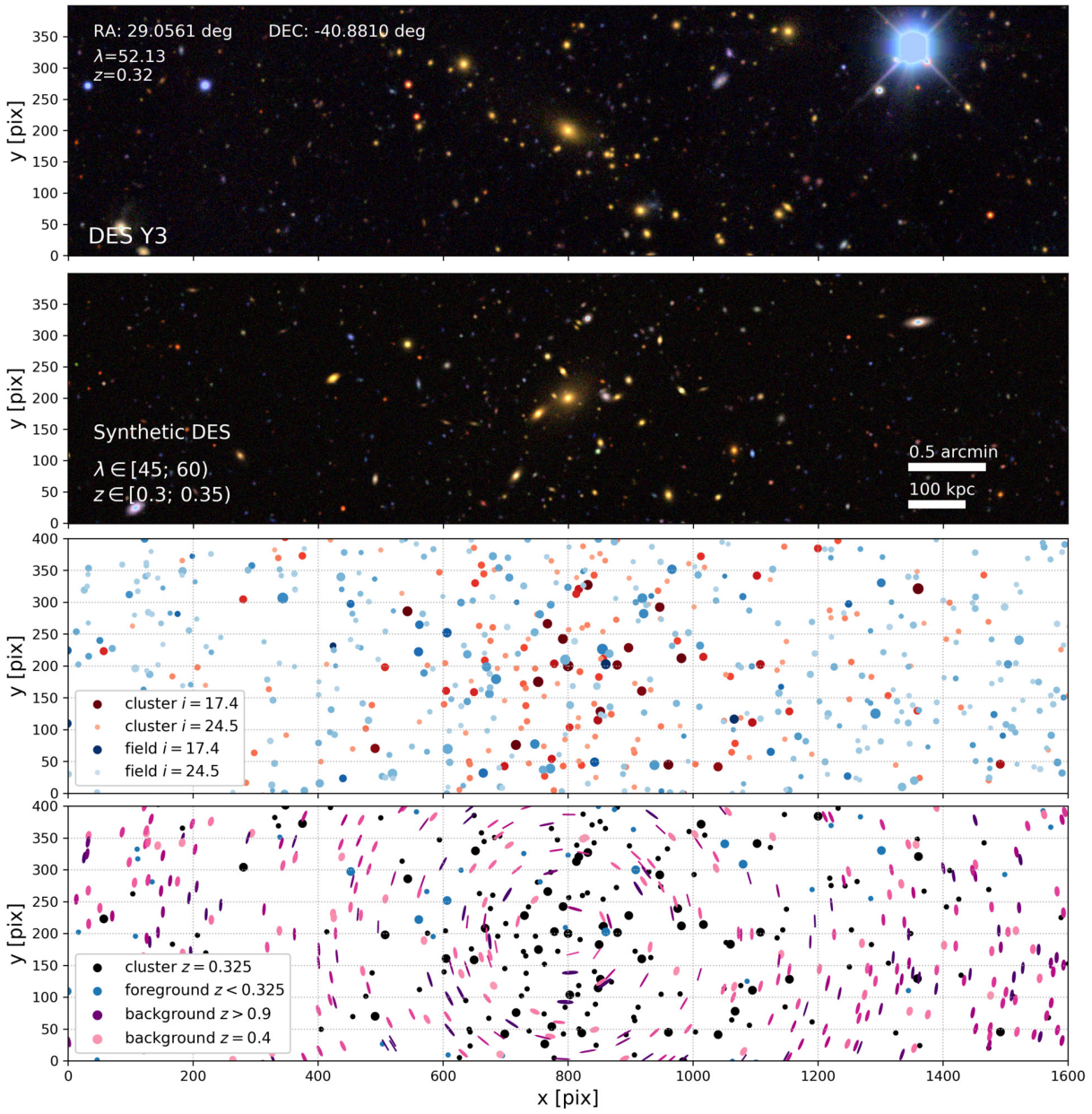

Figure 4.22: Real and synthetic galaxy cluster side by side. *Top: gri* color composit image of a real redMaPPer galaxy cluster in the DES Y3 footprint. *Second row: gri* color composit image of a synthetic galaxy galaxy cluster representative of $\lambda \in [45\ 60)$, $z \in [0.3; 0.35)$. *Third row:* brightness distribution of the synthetic light sources for cluster members (red/brown) and foreground and background objects (blue). Darker shades and larger symbols correspond to brighter objects. *Bottom row:* exaggerated shear map of background sources (magenta ellipses) with the shade representing redshift, cluster members (black) and foreground sources (blue).

An illustration of this joint BCG + ICL light profile in the mock cluster images is shown on Figure 4.21. The two panels show an identical set of mock galaxies for a synthetic cluster corresponding to the cluster bin with $\lambda \in [45; 60)$ and $z \in [0.3; 0.35)$, however the left panel shows only the `ngmix` galaxy models, while the right panel also shows the ICL component added using the above formula.

4.5.4 Synthetic Survey-like Images

The last step of creating mock observations is to render the above defined components into synthetic survey-like images. Simulated galaxy images are the bedrock of estimating the performance of weak lensing methods, and therefore they were the topic of extensive study in recent literature (Massey et al., 2007; Bridle et al., 2009; Mandelbaum et al., 2015). These efforts also extend to the calibration and validation of the DES weak lensing shape estimation methods (Jarvis et al., 2016; Zuntz et al., 2018; Samuroff et al., 2018). In the following we make use of a simplified version of the image simulation pipeline developed for the Y3 data release of DES (N. et al., in prep).

The construction of a synthetic observation starts with a catalog of photometric objects which will inhabit the mock image. For this study this catalog contains the parameters of the `ngmix` light distribution model for each entry which are pixel position in the image, shape $(g_1; g_2)$, size T, De Vaucouleours fraction, and fluxes in g, r, i, z bands. This catalog corresponds to a random realization of a mock line-of-sight constructed according to section 4.5.1. The galaxy shapes are defined according to section 4.5.2 where magnification effects are also applied. Finally the central galaxy is added as defined in section 4.5.3. At this stage stars and foreground objects can be added according to their density at the targeted galactic latitude of the mock observations. In the present pathfinder study we only consider a simplified scenario and add a stellar sample drawn from the deep field catalog according to their relative density in the deep field footprints.

Synthetic images are created via a customized version of the DES Y3 image simulation pipeline (N. et al., in prep), which renders images based on a galaxy image simulation package `GalSim` (Rowe et al., 2015), while using an extension package for the `ngmix` light profile model used in the actual DES Y3 analysis[6]. This model describes the galaxies as a combination of two terms: an exponential light profile (disk) and a De Vaucouleours (bulge) light profile. Given that most galaxies in a DES-like survey are poorly resolved, an additional constraint is enforced by setting the effective radius of both light profile components to be identical.

In the following, we consider a simplified setup of the observational scenario of DES where we directly simulate the so-called *co-added* survey images. Under real circumstances due to variations in observing conditions and the point spread function (PSF) between exposures the net PSF in co-added images is difficult to model, hence the DES shape estimation pipeline itself takes single exposure images as input. In a simulation such variations can be factored out and that allows us to simplify the simulation setup into deeper mock co-added images with well-behaving PSFs.

The synthetic co-added images are constructed the following way:

[6] `https://github.com/esheldon/ngmix`, the `ngmix.gmix.GmixBDF` model.

1. First the image canvas is defined with its desired dimensions and pixel scale, which in the case of DES is 0.27 arcsec / pixel. In the DES Y3 image simulation pipeline the canvas is defined as a 10k$\times$10k pixel rectangle, however in the current analysis only the inner 10 arcminute ($\approx$ 2222 pixel) radius circle will be populated with photometric objects.

2. For each object a small cutout image (postage stamp) is constructed. The light model is defined using `ngmix`, convolved with a representation of the mock PSF, then rendered into a postage stamp. The size of the postage stamp is defined to contain approximately 99% of the light of the galaxy. In each postage stamp the galaxy is approximately located in the center, with small offsets applied to account for the position of the light source relative to the pixel grid. In the case of the current pathfinder cluster simulations the PSF is modeled as a Gaussian with a full-width half-maximum (FWHM) of 0.9 arcsec, which is roughly equal to the median DES observing condition (Drlica-Wagner et al., 2018).

3. After the creation of all postage stamps from the input catalog, they are added onto the main canvas at their intended pixel positions.

4. A noise map is applied to the image. This can be either a homogeneous noise over the entirety of the image, or a localized noise map intended to mimic the depth variations and gaps of the underlying single epoch images. In the current set of mock clusters we take the noise properties of a randomly selected DES tile (DES2122+0209) and apply Gaussian noise matched to reproduce the median flux of the unmasked regions of the reference tile in the chosen observational band. Choosing the noise level for synthetic images is however not straightforward, as a substantial amount of light which is traditionally attributed to the sky level or to background noise in fact originates from undetected faint stars and galaxies (Hoekstra et al., 2017; Euclid Collaboration et al., 2019; Eckert et al., 2020). In the framework of the present analysis many of these undetected sources are explicitly part of the rendered objects, therefore as a rough approximation we reduce the background noise variance by half for illustration purposes. In a more targeted analysis the noise level reduction should be set iteratively such that the combination of artifical noise and undetected sources reproduces the measured noise.

5. Finally the tapered ICL model defined according to section 4.5.3 is evaluated for the pixel positions of the mock image and the additional light component is added onto the synthetic observation. In this step to ensure the light profile and the BCG joins continuously, we assume that the ICL has the same ellipticity and major axis direction alignment as the central galaxy.

This concludes the creation of a single mock observation, and the steps are repeated for each simulated observational band.

The end result of the above algorithm is a synthetic image of mock galaxy clusters constructed according to their measured optical properties in DES observations. The composition of this type of mock cluster is illustrated on Figure 4.3 and Figure 4.22 where a *gri*-band color composit image is shown for synthetic clusters side by side with redMaPPer clusters with similar

observable parameters. The real and simulated survey images appear very similar qualitatively, with the obvious difference of observational imperfections and masked objects appearing in the real observations. While the synthetic images do contain an approximate stellar population based on faint stars observed in the deep field, very bright stars which need to be masked or low redshift foreground galaxies are not currently reproduced in the mock observations. Furthermore low redshift foreground objects such as galaxies with visible disc and spiral arm features are not contained in the scope of the present analysis.

In addition to the color composite images Figure 4.3 and Figure 4.22 also illustrate the composition of the lines-of-sight. The third row of each figure shows the brightness distribution of the cluster component with brown/red symbols, and the foreground and background component with blue symbols. The shade and size of the symbols indicate the brightness with fainter objects shown as smaller markers. It is visible that many of the faint objects are barely or not at all discernible on the composite images. Yet these unresolved sources influence the performance of photometric methods (Hoekstra et al., 2017; Euclid Collaboration et al., 2019). The bottom row of each figure shows the exaggerated gravitational shear imprinted on background sources (the ellipticities are increased by a factor of 20). The background sources are shown in as darker color for low redshift and lighter color for high redshifts. Cluster members are shown in black symbols with, while foreground objects are shown as blue. The different brightness values are indicated by their different sizes.

While the galaxy populations of the $\lambda \in [30; 45)$ and $\lambda \in [45; 60)$ bins are found to be close in terms of their galaxy surface density profiles, clusters show greater differences between the different redshift ranges. This is illustrated by Figure 4.23, which shows synthetic galaxy clusters with $\lambda \in [45; 60)$ in the $z \in [0.3; 0.35)$, $z \in [0.45; 0.5)$ and $z \in [0.6; 0.65)$ cluster samples. These color composit images show a striking illustration of the changes in the visible properties of galaxy clusters across cosmic time. Additional examples of mock clusters for the different redshift ranges are shown on Figure 4.24, Figure 4.25 and Figure 4.26.

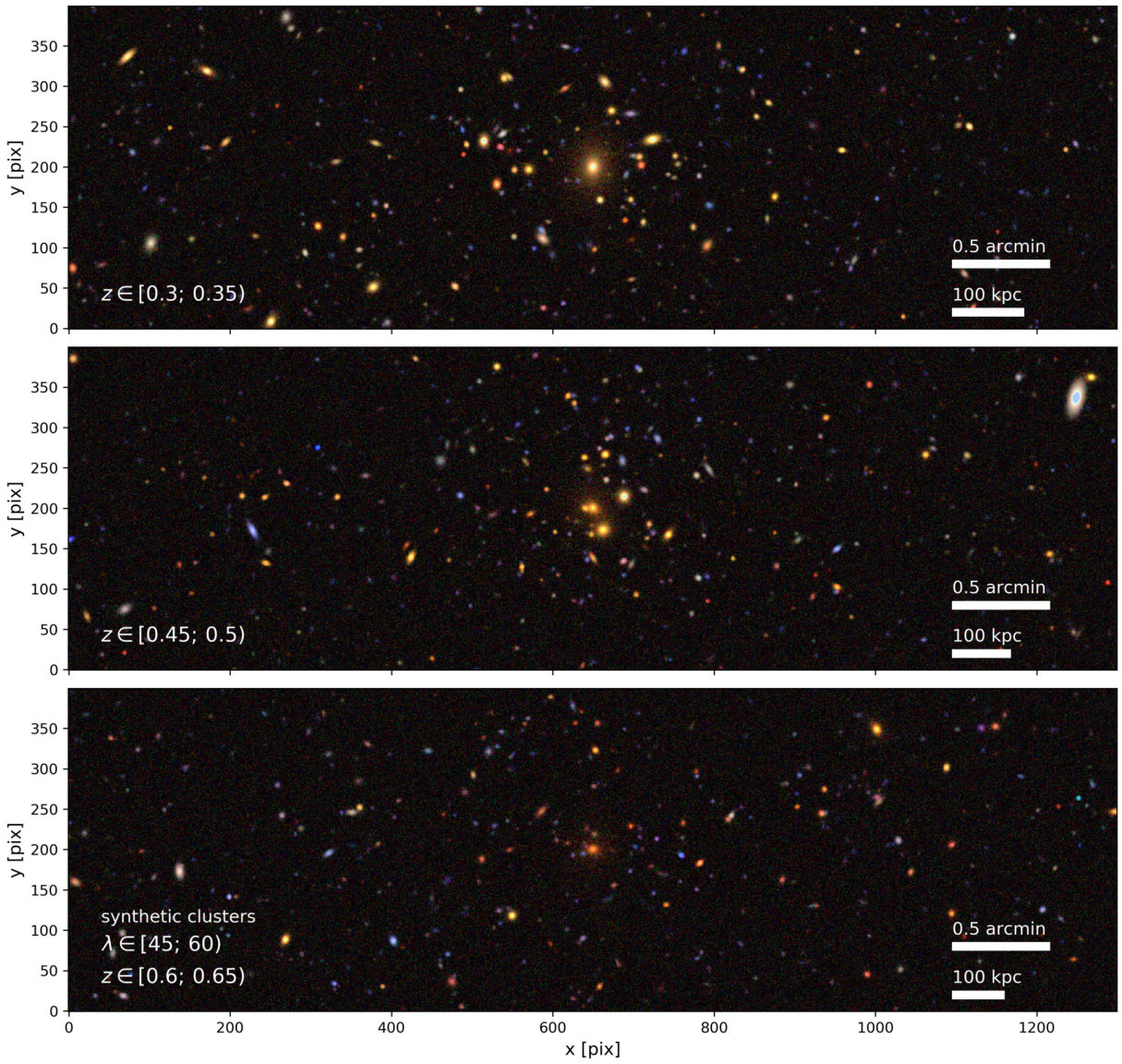

Figure 4.23: Synthetic galaxy clusters corresponding to redMaPPer clusters with $\lambda \in [45;\ 60)$ across the different redshift ranges

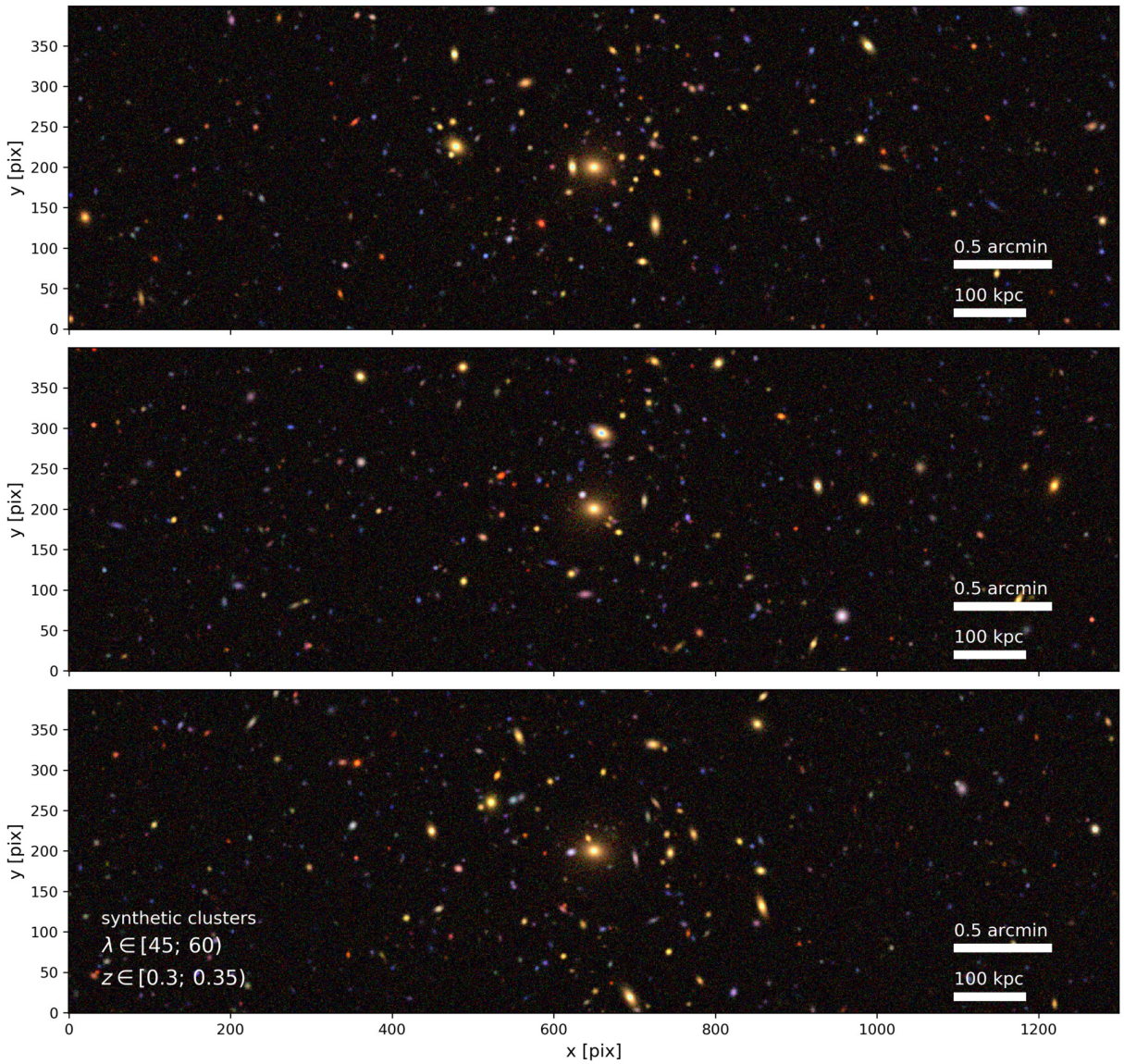

Figure 4.24: Synthetic galaxy clusters corresponding to redMaPPer clusters with $z \in [0.3;\ 0.35)$ and $\lambda \in [45;\ 60)$

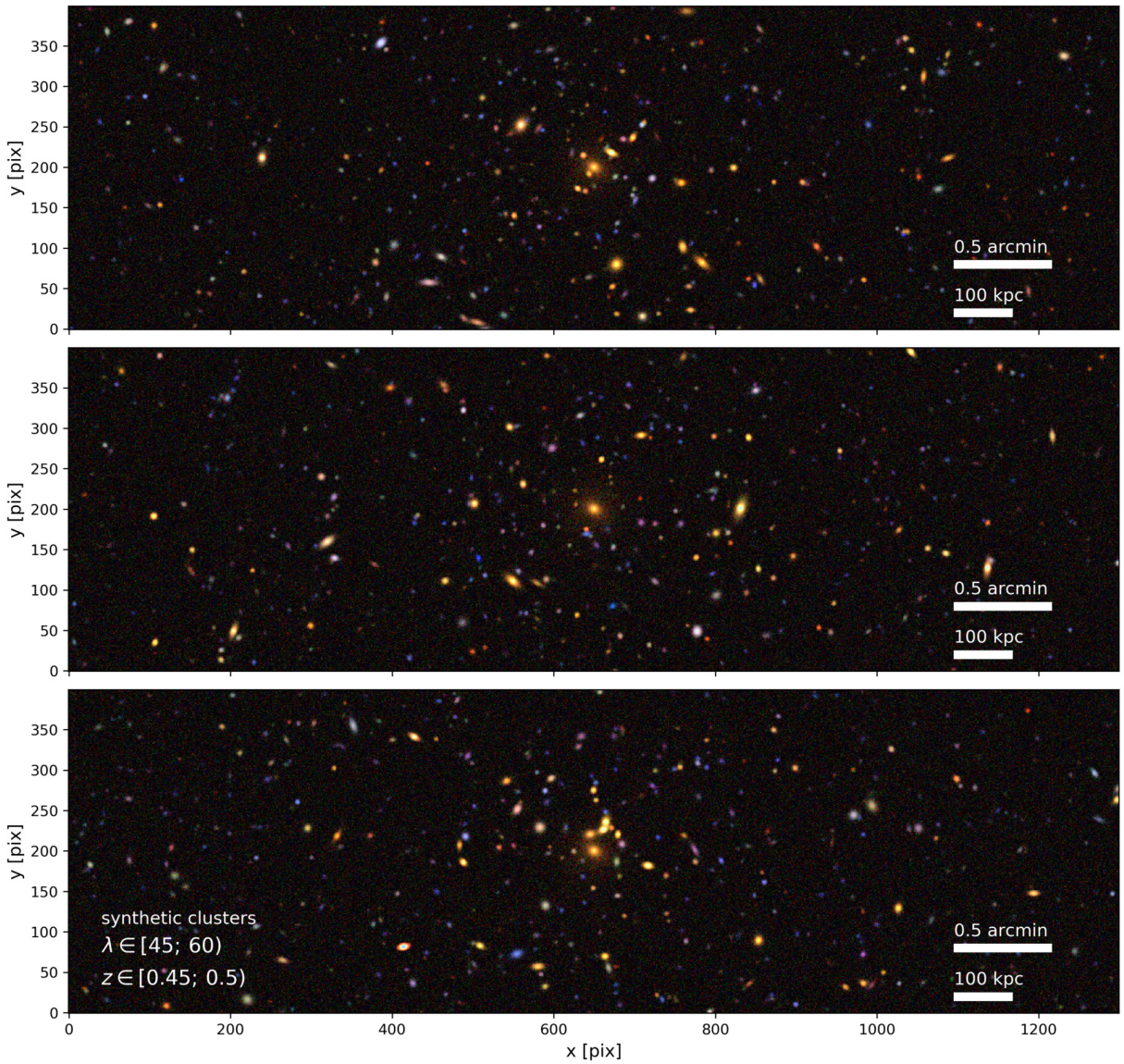

Figure 4.25: Synthetic galaxy clusters corresponding to redMaPPer clusters with $z \in [0.45;\ 0.5)$ and $\lambda \in [45;\ 60)$

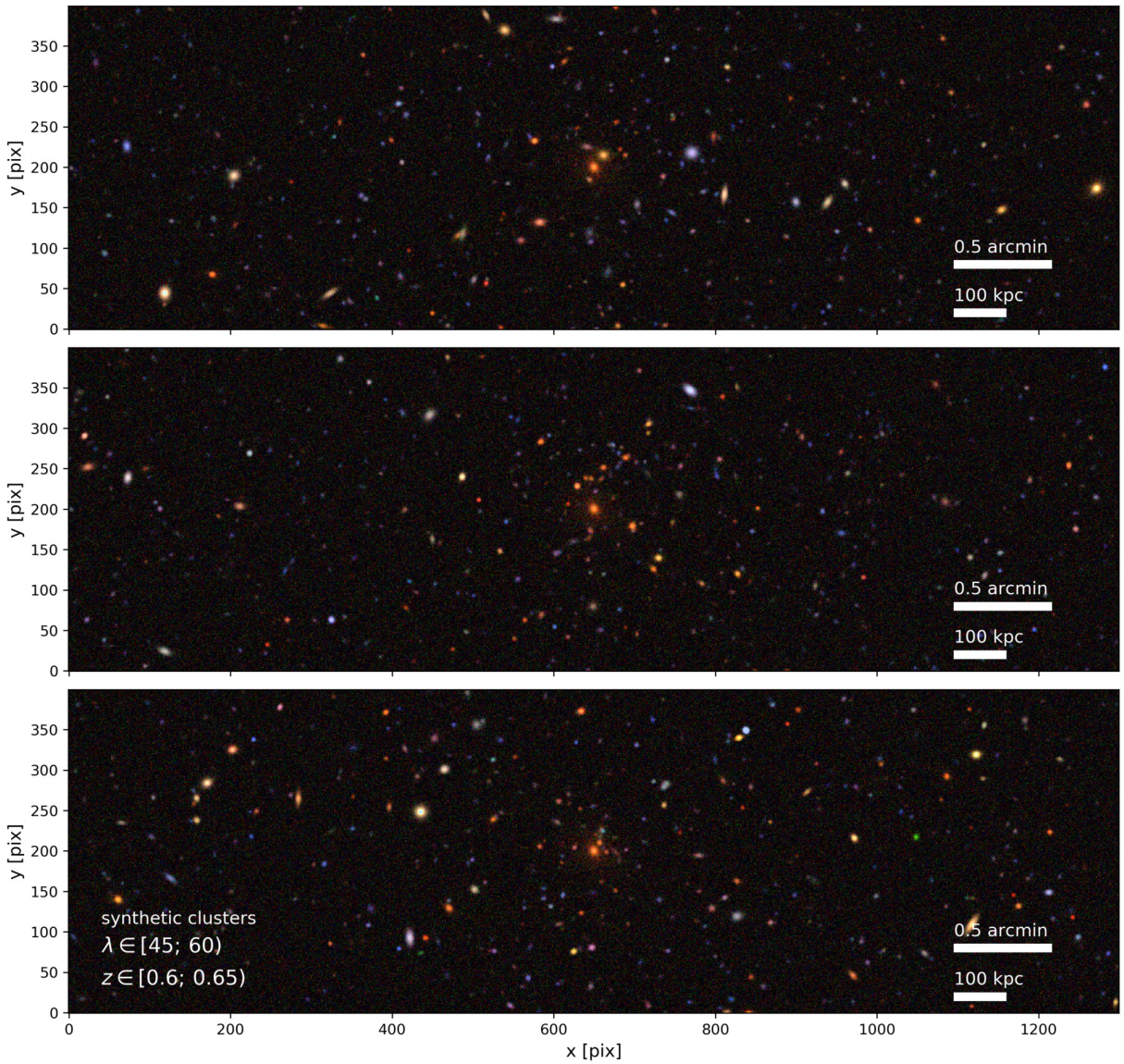

Figure 4.26: Synthetic galaxy clusters corresponding to redMaPPer clusters with $z \in [0.6;\ 0.65)$ and $\lambda \in [45;\ 60)$

4.6 Summary and Conclusions

We have presented the results and algorithm of a pathfinder study aiming to generate synthetic galaxy cluster observations purely from observational data taken by the Dark Energy Survey up to its third year of observations. The focus of this work was to construct a test environment for galaxy cluster weak lensing studies where the interplay between analysis choices and different types of systematic effects related to source galaxy selection, shape and photometric redshift estimation can be studied in a controlled manner.

In the presented analysis we

- Measured the galaxy feature correlation function around clusters in projection (section 4.3.1).
- Developed a framework for representing galaxy distributions as p.d.f.-s via KDE (section 4.3.2).
- Performed a statistical background subtraction (section 4.3.3.1) and (section 4.3.3.3).
- Extrapolated survey depth using information from the deep fields (section 4.3.3.2).
- Combined the above steps into a recipe on how to render mock galaxy clusters (section 4.5).

The statistical results of this work are presented in section 4.4, while example realisations of synthetic galaxy cluster observations are shown on Figure 4.23, Figure 4.24, Figure 4.25 and Figure 4.26.

Our method addresses three distinct problems arising with simulated data:

A Synthetic galaxy clusters are generated to match their observed properties in the Dark Energy Survey, and extrapolations of the galaxy populations are performed where necessary, based also directly on observational data from the deep fields of DES. This approach ensures that the created synthetic observations match observed reality in terms of their galaxy populations. Therefore the method does not rely on numerical simulations of baryonic structure formation and galaxy evolution to construct galaxy clusters and hence it is independent from assumptions and approximations inherent in cosmological simulations.

B Via the statistical learning approach new realizations of synthetic observations can be created at minimal computational cost. This allows the creation of large samples of mock cluster images which in turn allow the precise and accurate characterization of systematic uncertainties in cluster lensing measurements.

C The used algorithm is formulated as a explicit recipe instead of a black box model such as a neural network, therefore the different components can be readily modified if need be and external information can be added in a targeted way.

This work was done in preparation of a dedicated cluster weak lensing benchmark analysis for the Dark Energy Survey, and the details of the algorithm were tuned to the format of that dataset. However the method is easily generalized to other similar large area sky surveys such as LSST (LSST Science Collaboration et al., 2009) and Euclid (Amendola et al., 2018) as well.

Chapter 5

Summary and Conclusions

Weak gravitational lensing provides a powerful and fundamental way to connect optical observations with the properties of the hidden dark sector of the Universe. In order to make use of this mode of inference, in this thesis we developed the analysis formalism and measured, calibrated and characterized weak gravitational lensing profiles and halo masses within the observational dataset of the Dark Energy Survey (DES).

Our first aim was to perform the weak lensing mass calibration of galaxy clusters selected by the redMaPPer algorithm in the DES year 1 data, which serves as the input of the DES cluster cosmology analysis (McClintock & Varga et al., 2019). As part of this work:

A1 We developed the formalism for weak lensing profile measurements and calibrations in the DES Y1 data. We split clusters into 4×3 bins of richness λ and redshift z for $\lambda \geq 20$ and $0.2 \leq z \leq 0.65$ and measured their respective stacked weak lensing signal.

A2 We calculated a semi-analytic covariance matrix to characterize the statistical errors in the recovered weak lensing profiles.

A3 We derived the mean cluster masses for the richness and redshift bins while accounting for the different calibrations and systematic uncertainties at the likelihood level.

A4 Our analysis accounted for the following sources of systematic error: shear and photometric redshift errors, cluster miscentering, cluster member dilution of the source sample, systematic uncertainties in the modeling of the halo–mass correlation function, halo triaxiality, and projection effects.

A5 We modeled the richness - redshift - mass relation and constrained the normalization of the scaling relation at the 5.0 per cent level. This is the tightest measurements made to date, and our result is dominated by systematic uncertainities which alone reach 4.3 per cent. Alas, the calibration of the richness - redshift - mass relation still remains the limiting factor for testing cosmological models with cluster counts.

As part of the above weak lensing mass calibration, we carried out a detailed validation of a method for correcting the lensing signal for cluster members contaminating the source galaxy catalog and diluting the lensing signal (Varga et al., 2019). During this work:

B1 We decompose the photometric redshift $P(z)$ of source galaxies used in the weak lensing measurement into contaminating and background components, and estimate the contamination from the relative amplitude of these two components.

B2 We tested the method in a mock survey simulation, finding excellent agreement between the estimated contamination rates and the actual true number of contaminants extracted from the truth catalogs of the simulation. From this comparison we estimate a systematic uncertainty of $\delta_{sys} < 1$ for the $P(z)$ decomposition method.

B3 We investigated the validity of the Gaussian ansatz for the cluster $P(z)$ component within the mock analysis. We found that this method recovers the approximate redshift and width of the peak within the $P(z)$ of the contaminating galaxies.

B4 We tested the sensitivity of the contamination estimates to the choice of the background $P(z)$ component on DES Y1 data, and found an excellent agreement between the boost factors derived via the fiducial and alternative background components.

B5 We compared the method with an alternative, transverse correlation based contamination estimate. We found indications that the alternative method is preferentially underestimating the contaminations, which is likely an imprint of the radial source galaxy selection function. This is expected to impact the decomposition predictions to a lesser extent as it does not make use of the number profile of sources.

Finally, motivated by the utility of the mock survey simulation used in the previous step, we construct synthetic optical observations of galaxy clusters to create a test environment for cluster weak lensing studies, where the interplay between analysis choices and different types of systematic effects related to source galaxy selection, shape, and photometric redshift estimation can be studied in a controlled manner (Varga et al., in prep). As part of this work:

C1 We measured the projected galaxy distribution around redMaPPer selected galaxy clusters in the DES year 3 data in two bins of cluster richness ($\lambda \in [30; 45)$ and $\lambda \in [45; 60)$) and three narrow bins in cluster redshift ($z \in [0.3; 0.35)$, $z \in [0.45; 0.5)$ and $z \in [0.6; 0.65)$).

C2 We developed a statistical learning framework for representing the radial distribution of galaxy features (colors, magnitudes, morphology parameters) around the galaxy cluster as probability distribution functions.

C3 We estimated the feature distribution of cluster member galaxies by comparing the galaxy feature distributions around galaxy clusters and the corresponding set of reference random points. In contrast to redMaPPer our formulation also includes non-red-sequence and faint cluster member galaxies.

C4 We extrapolated the cluster member galaxy feature distribution below the limiting magnitude of the wide field DES survey by incorporating information about the galaxy distribution in the DES deep fields.

C5 From the combination of the above steps we defined a recipe for constructing mock galaxy clusters as random draws from the feature probability distribution function of cluster member galaxies.

C6 We rendered mock cluster and line-of-sight catalogs into synthetic cluster images in the same format as actual survey observations.

In this framework the synthetic galaxy clusters are thus generated to match the galaxy distributions in real observational data, and they are thus independent from the assumptions and approximations inherent in cosmological simulations. Furthermore new realizations of synthetic cluster images can be created at minimal computational cost, and in turn this will allow future analyses the precise and accurate characterization of systematic uncertainties in cluster lensing measurements based on large suites of synthetic cluster images.

Impact on the Scientific Field

The weak lensing profile measurement setup and the cluster member contamination calibration used for the DES year 1 (Y1) analysis was collated in the `xpipe` analysis pipeline designed and implemented by myself. The measurement pipeline is now part of DES infrastructure and is documented and made publicly available[1].

The DES Y1 galaxy cluster mass calibration has also enabled a series of further scientific studies, whose results we discuss briefly in the following:

- Chang et al. (2018b) investigated the so-called splashback feature around galaxy clusters in DES. The splashback feature relates to the first apocenter, or turnaround point of matter accreting onto the dark matter halo, and the splashback radius r_{sp} defines a boundary for the halo which is connected to the dynamics of the cluster, as opposed to simply the overdensity. In simple terms, the splashback feature manifests as a rapid steepening in the halo mass profile near r_{sp}. In their analysis both the galaxy number density profiles and the weak lensing mass profiles were measure around a sample of redMaPPer selected galaxy clusters in DES Y1 with $\lambda \in [20;\ 100)$ and $z \in [0.2;\ 0.55)$, and the weak lensing profile measurement was performed by myself based on the formalism and contamination correction presented in this thesis. The mean mass for the cluster sample was estimated to be $M_{200m} = 2.5 \times 10^{14}\ \mathrm{M}_{\odot}$, and a strong evidence of a splashback-like steepening of the galaxy density profile is found at $r_{sp}^{\mathrm{gal}} = 1.13 \pm 0.07$ Mpc/h. Furthermore the study demonstrates for the first time the existence of a splashback-like steepening of the matter profile of galaxy clusters derived purely from weak lensing measurements with $r_{sp}^{\mathrm{WL}} = 1.34 \pm 0.21$ Mpc/h.

[1] `https://xpipe.readthedocs.io`

- Pereira et al. (2020) presents a mass calibration of an alternative cluster mass proxy: $\mu_\star$ derived from the stellar masses of galaxy clusters in DES Y1. Stellar mass is expected to have less intrinsic scatter with respect to halo mass than richness (e.g. the λ defined by redMaPPer) and is furthermore estimated to be less affected by projection effects. Hence this kind of alternative mass calibration is an important eventual consistency check for cosmology analyses based on the masses and number counts of galaxy clusters. In their calibration they use $\sim 6,000$ galaxy clusters initially identified by redMaPPer , but selected based on their estimated stellar masses and redshifts into 12 subsets within $0.1 \leq z < 0.65$ and stellar mass ($\mu_\star$) up to $\sim 5.5 \times 10^{13}$ $M_\odot$. Their analysis follows closely that of the work presented in this thesis by fitting a joint mass–$\mu_\star$–z scaling relation using cluster masses derived from the stacked weak lensing signal in the individual bins, and their weak lensing profiles were measured by myself using the `xpipe` package developed in this thesis.

- DES Collaboration et al. (2020) presents the Y1 cluster cosmology results of DES based on a joint analysis of the counts and weak lensing masses of redMaPPer clusters, using the weak lensing mass calibration described in this thesis. They report surprisingly low values for $S_8 = \sigma_8(\Omega_m/0.3)^{0.5} = 0.65 \pm 0.04$ and $\Omega_m = 0.179^{+0.031}_{-0.038}$. These $\sigma_8 - \Omega_m$ posteriors are in 2.4σ tension with the DES Y1 3x2pt results (Abbott et al., 2018), and in 5.6σ tension with the Planck CMB analysis (Planck Collaboration et al., 2018). Internal consistency tests, and the consistency of the cluster number counts with external data indicate that the tension is primarily driven by the mass estimates of low richness ($\lambda < 30$) clusters. Given that systematic effects of the weak lensing measurement itself impact this richness range the least, the cause of the tension is likely to be the incorrect interpretation of the lensing profiles when deriving the cluster masses. This can be expressed as the impact of the photometric cluster selection on the stacked lensing profiles of clusters, which results in the derived cluster masses to be biased low compared to a mass selected cluster sample. As a selection effect which could give rise to the necessary level of mass bias was not yet identified, the better understanding of the details of photometric cluster selection from different observational and simulation based efforts have become the priority for the following era of research.

- The *lensing without borders* (LWB; Leauthaud et al., in prep) is an ongoing comparison project aiming to test the consistency of weak lensing measurements across six distinct wide area lensing surveys: SDSS, CS82, CFHTLenS, DES, HSC, and KiDS. The main idea of this project is that since $\Delta\Sigma$ is a physical quantity, the measured values for a shared sample of lens galaxies should agree between the analysis choices and approaches of the different surveys. The goal of LWB is to perform a blind comparison of the $\Delta\Sigma$ amplitudes, and in turn this provides an empirical end-to-end test of systematics in weak lensing measurements. The DES side of these calculations were performed by myself using the `xpipe` package developed in this thesis. While this work is still under preparation, the initial post un-blinding results seem to indicate an overall broad consistency between the measurements of the different surveys and DES. However this result is limited by the rela-

tively large statistical uncertainty caused by the typically small overlap region of the DES Y1 footprint and the footprint of the reference set of lens galaxies.

Closing Remarks

- In this thesis we measured, calibrated and characterized the $\Delta\Sigma$ profiles of galaxy clusters in the year 1 dataset of the Dark Energy Survey. These results passed an extensive set of tests conducted both in this thesis and in the other scientific studies enabled by this work.
- The observed tension in the DES cluster cosmology analysis likely relates to the interpretation of the lensing profiles, and not to a bias in the lensing profiles themselves. For this reason future studies must focus on the multi-probe and multi-wavelength characterization and calibration of the selection effects in galaxy cluster catalogs defined based on optical data.
- With the increasing statistical power, and improved calibration of weak lensing measurements new and innovative science cases relating to the non-linear evolution of structure can be studied with current and upcoming lensing datasets. In addition to the splashback feature, the mass evolution of cluster member galaxies also presents a lucrative direction, which however necessitates the the creation of dedicated end-to-end simulations of galaxy cluster weak lensing measurements in a controlled synthetic environment.

Appendix A

WL Mass Calibration of redMaPPer Galaxy Clusters

A.1 Acknowledgements of McClintock, Varga et al. (2019)

This paper has gone through internal review by the DES collaboration. TM and ER are supported by DOE grant DE-SC0015975. TNV and SS are supported by the SFB-Transregio 33 'The Dark Universe' by the Deutsche Forschungsgemeinschaft (DFG) and the DFG cluster of excellence 'Origin and Structure of the Universe'. ER acknowledges additional support by the Sloan Foundation, grant FG-2016-6443. Support for DG was provided by NASA through Einstein Postdoctoral Fellowship grant number PF5-160138 awarded by the Chandra X-ray Center, which is operated by the Smithsonian Astrophysical Observatory for NASA under contract NAS8-03060. TJ, DH, and SE are supported by DOE grant DE-SC0010107. KR and SB acknowledge support from the UK Science and Technology Facilities Council via Research Training Grants ST/P000252/1 and ST/N504452/1, respectively. CV is supported by the Mexican National Council for Science and Technology via grant number 411117.

Part of our computations have been carried out on the computing facilities of the Computational Center for Particle and Astrophysics (C2PAP). A portion of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da

Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2015-71825, ESP2015-66861, FPA2015-68048, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through project number CE110001020, and the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) e-Universe (CNPq grant 465376/2014-2).

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

This work used simulations and computations performed using computational resources at SLAC and at NERSC.

A.2 The redMaPPer v6.4.17 cluster catalog

The full redMaPPer DES Y1A1 catalogs will be available at `http://risa.stanford.edu/redmapper/` in FITS format. The catalogs will also be available from the online journal in machine-readable formats. We note that this is of the same

format as (Rykoff et al., 2016), and we point the reader to that paper for further details. The cluster catalog is described in Table A.1, and the associated members in Table A.2. The catalog is the "full" catalog, with all clusters with $\lambda > 20$, and the volume-limited subset is flagged with the `VLIM` flag. The map of the maximum redshift of the volume-limited catalog is described in Table A.3, and the random points are described in Table A.4.

A.3 Parameter posteriors

When fitting the weak lensing profiles some parameters are not constrained by a prior and are also not shared between cluster bins. These are the halo concentration c, the boost factor amplitude B_0, and the boost factor scale radius R_s. Table A.5 shows the posteriors for these three parameters for each cluster bin. As seen in Figure 2.10 B_0 and R_s are highly degenerate.

Table A.1: redMaPPer Y1A1 Cluster Catalog Format

Name	Format	Description
ID	INT(4)	redMaPPer Cluster Identification Number
VLIM	INT(2)	One if in cosmology catalog, 0 otherwise
NAME	CHAR(20)	redMaPPer Cluster Name
RA	FLOAT(8)	Right ascension in decimal degrees (J2000)
DEC	FLOAT(8)	Declination in decimal degrees (J2000)
Z_LAMBDA	FLOAT(4)	Cluster photo-z z_λ
Z_LAMBDA_ERR	FLOAT(4)	Gaussian error estimate for z_λ
LAMBDA	FLOAT(4)	Richness estimate λ
LAMBDA_ERR	FLOAT(4)	Gaussian error estimate for z_λ
S	FLOAT(4)	Richness scale factor
Z_SPEC	FLOAT(4)	Spectroscopic redshift for most likely center (-1.0 if not available)
COADD_OBJECTS_ID	INT(8)	DES `COADD_OBJECTS_ID` identification number
MAG_CM_G	FLOAT(4)	g MAG_CM magnitude for most likely central galaxy (SLR corrected)
MAGERR_CM_G	FLOAT(4)	error on g MAG_CM magnitude
MAG_CM_R	FLOAT(4)	r MAG_CM magnitude for most likely central galaxy (SLR corrected)
MAGERR_CM_R	FLOAT(4)	error on g MAG_CM magnitude
MAG_CM_I	FLOAT(4)	i MAG_CM magnitude for most likely central galaxy (SLR corrected)
MAGERR_CM_I	FLOAT(4)	error on g MAG_CM magnitude
MAG_CM_Z	FLOAT(4)	z MAG_CM magnitude for most likely central galaxy (SLR corrected)
MAGERR_CM_Z	FLOAT(4)	error on g MAG_CM magnitude
ZLUM	FLOAT(4)	Total membership-weighted z-band luminosity (units of L_*)
P_CEN[5]	5×FLOAT(4)	Centering probability $P_{\rm cen}$ for 5 most likely centrals
RA_CEN[5]	5×FLOAT(8)	R.A. for 5 most likely centrals
DEC_CEN[5]	5×FLOAT(8)	Decl. for 5 most likely centrals
ID_CEN[5]	5×INT(8)	DES `COADD_OBJECTS_ID` identification number for 5 most likely centrals
PZBINS[21]	21×FLOAT(4)	Redshift points at which $P(z)$ is evaluated
PZ[21]	21×FLOAT(4)	$P(z)$ evaluated at redshift points given by PZBINS

Table A.2: redMaPPer DES Y1A1 Member Catalog Format

Name	Format	Description
ID	INT(4)	redMaPPer Cluster Identification Number
RA	FLOAT(8)	Right ascension in decimal degrees (J2000)
DEC	FLOAT(8)	Declination in decimal degrees (J2000)
R	FLOAT(4)	Distance from cluster center (h^{-1} Mpc)
P	FLOAT(4)	Membership probability
P_FREE	FLOAT(4)	Probability that member is not a member of a higher ranked cluster
THETA_L	FLOAT(4)	Luminosity (z-band) weight
THETA_R	FLOAT(4)	Radial weight
MAG_CM_G	FLOAT(4)	g MAG_CM magnitude (SLR corrected)
MAGERR_CM_G	FLOAT(4)	error on g MAG_CM magnitude
MAG_CM_R	FLOAT(4)	r MAG_CM magnitude (SLR corrected)
MAGERR_CM_R	FLOAT(4)	error on r MAG_CM magnitude
MAG_CM_I	FLOAT(4)	i MAG_CM magnitude (SLR corrected)
MAGERR_CM_I	FLOAT(4)	error on i MAG_CM magnitude
MAG_CM_Z	FLOAT(4)	z MAG_CM magnitude (SLR corrected)
MAGERR_CM_Z	FLOAT(4)	error on z MAG_CM magnitude
Z_SPEC	FLOAT(4)	Spectroscopic redshift (-1.0 if not available)
COADD_OBJECTS_ID	INT(8)	DES `COADD_OBJECTS_ID` identification number

Table A.3: redMaPPer $z_{\rm max}$ Map Format

Name	Format	Description
HPIX	INT(8)	`HEALPIX` ring-ordered pixel number (`NSIDE=4096`)
ZMAX	FLOAT(4)	Maximum redshift of a cluster centered in this pixel
FRACGOOD	FLOAT(4)	Fraction of pixel area that is not masked

Table A.4: redMaPPer Random Points Catalog Format

Name	Format	Description
RA	FLOAT(8)	Right ascension in decimal degrees (J2000)
DEC	FLOAT(8)	Declination in decimal degrees (J2000)
Z	FLOAT(4)	Redshift of random point
LAMBDA	FLOAT(4)	Richness of random point
WEIGHT	FLOAT(4)	Weight of random point

Table A.5: Lensing profile parameters not constrained by priors or shared between cluster bins. Uncertainties are the 68 per cent confidence intervals. Note that in the highest redshift and richness bin, the boost factor profile model scale radius had a bimodal distribution. This did not affect the mass estimate at all.

λ	$z \in [0.2, 0.35)$	$z \in [0.35, 0.5)$	$z \in [0.5, 0.65)$
Concentration c			
$[20, 30)$	5.81 ± 1.03	5.68 ± 1.14	4.76 ± 1.62
$[30, 45)$	4.53 ± 0.74	6.24 ± 1.08	3.61 ± 0.72
$[45, 60)$	4.38 ± 0.96	5.41 ± 1.17	4.76 ± 1.21
$[60, \infty)$	4.65 ± 0.82	3.19 ± 0.56	3.73 ± 1.02
Boost factor amplitude B_0			
$[20, 30)$	0.34 ± 0.05	0.05 ± 0.01	0.13 ± 0.05
$[30, 45)$	0.37 ± 0.06	0.14 ± 0.04	0.13 ± 0.08
$[45, 60)$	0.27 ± 0.05	0.05 ± 0.02	0.09 ± 0.06
$[60, \infty)$	0.23 ± 0.03	0.21 ± 0.17	0.03 ± 0.04
Boost factor scale radius R_s [Mpc]			
$[20, 30)$	0.44 ± 0.06	0.89 ± 0.24	0.38 ± 0.11
$[30, 45)$	0.50 ± 0.07	0.44 ± 0.10	0.44 ± 0.18
$[45, 60)$	0.80 ± 0.15	1.72 ± 0.95	0.85 ± 0.37
$[60, \infty)$	1.37 ± 0.21	0.51 ± 0.23	35.94 ± 29.69

Appendix B

Validation of weak lensing cluster contamination estimates

B.1 Acknowledgements of Varga et al. (2019)

This paper has gone through internal review by the DES collaboration. TNV and SS are supported by the SFB-Transregio 33 'The Dark Universe' by the Deutsche Forschungsgemeinschaft (DFG) and the DFG cluster of excellence 'Origin and Structure of the Universe'. JD, DG and RHW received support from the U.S. Department of Energy under contract number DE-AC02-76SF00515. DG was also supported by Chandra Award Number GO8-19101A, issued by the Chandra X-ray Observatory Center. ER was supported by the DOE grant DE-SC0015975, by the Sloan Foundation, grant FG-2016-6443, and the Cottrell Scholar program of the Research Corporation for Science Advancement. AvdL is supported by the U.S. Department of Energy under Award Number DE-SC0018053.

Part of our computations have been carried out on the computing facilities of the Computational Center for Particle and Astrophysics (C2PAP). A portion of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. This research made use of computational resources at SLAC National Accelerator Laboratory, and the authors thank the SLAC computational team for support. Some of the computing for this project was performed on the Sherlock cluster. We would like to thank Stanford University and the Stanford Research Computing Center for providing computational resources and support that contributed to these research results. This research also used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois

at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2015-71825, ESP2015-66861, FPA2015-68048, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through project number CE110001020, and the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) e-Universe (CNPq grant 465376/2014-2).

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

Bibliography

Abbott, T. M. C., et al. *Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing*. *Phys. Rev. D*, 98(4):043526 (August 2018). 1708.01530.

Abbott, T. M. C., et al. *First Cosmology Results using Type Ia Supernovae from the Dark Energy Survey: Constraints on Cosmological Parameters*. *ApJ*, 872(2):L30 (February 2019). 1811.02374.

Aihara, H., et al. *The Hyper Suprime-Cam SSP Survey: Overview and survey design*. *PASJ*, 70:S4 (January 2018). 1704.05858.

Alam, S., et al. *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample*. *MNRAS*, 470(3):2617–2652 (September 2017). 1607.03155.

Allen, S. W., Evrard, A. E., & Mantz, A. B. *Cosmological Parameters from Observations of Galaxy Clusters*. *ARA&A*, 49:409–470 (September 2011). 1103.4829.

Amendola, L., et al. *Cosmology and fundamental physics with the Euclid satellite*. *Living Reviews in Relativity*, 21(1):2 (April 2018). 1606.00180.

Angulo, R. E. et al. *Scaling relations for galaxy clusters in the Millennium-XXL simulation*. *MNRAS*, 426:2046–2062 (November 2012). 1203.3216.

Annis, J., et al. *The Sloan Digital Sky Survey Coadd: 275 deg2 of Deep Sloan Digital Sky Survey Imaging on Stripe 82*. *The Astrophysical Journal*, 794(2):120 (2014).

Applegate, D. E., von der Linden, A., Kelly, P. L., Allen, M. T., Allen, S. W., Burchat, P. R., Burke, D. L., Ebeling, H., Mantz, A., & Morris, R. G. *Weighing the Giants - III. Methods and measurements of accurate galaxy cluster weak-lensing masses*. *MNRAS*, 439:48–72 (March 2014a). 1208.0605.

Applegate, D. E., von der Linden, A., Kelly, P. L., Allen, M. T., Allen, S. W., Burchat, P. R., Burke, D. L., Ebeling, H., Mantz, A., & Morris, R. G. *Weighing the Giants - III. Methods and*

measurements of accurate galaxy cluster weak-lensing masses. *MNRAS*, 439:48–72 (March 2014b). 1208.0605.

Bartelmann, M. & Schneider, P. *Weak gravitational lensing*. *Phys. Rep.*, 340:291–472 (January 2001).

Battaglia, N., et al. *Weak-lensing mass calibration of the Atacama Cosmology Telescope equatorial Sunyaev-Zeldovich cluster sample with the Canada-France-Hawaii telescope stripe 82 survey*. *J. Cosmology Astropart. Phys.*, 8:013 (August 2016). 1509.08930.

Baxter, E. J., Rozo, E., Jain, B., Rykoff, E., & Wechsler, R. H. *Constraining the mass-richness relationship of redMaPPer clusters with angular clustering*. *MNRAS*, 463(1):205–221 (November 2016). 1604.00048.

Baxter, E. J., et al. *A measurement of CMB cluster lensing with SPT and DES year 1 data*. *MNRAS*, 476(2):2674–2688 (May 2018). 1708.01360.

Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. *The ROCKSTAR Phase-space Temporal Halo Finder and the Velocity Offsets of Cluster Cores*. *ApJ*, 762:109 (January 2013). 1110.4372.

Benítez, N. *Bayesian Photometric Redshift Estimation*. *ApJ*, 536:571–583 (June 2000). astro-ph/9811189.

Bernstein, G. M. & Armstrong, R. *Bayesian lensing shear measurement*. *MNRAS*, 438:1880–1893 (February 2014a). 1304.1843.

Bernstein, G. M. & Armstrong, R. *Bayesian lensing shear measurement*. *MNRAS*, 438:1880–1893 (February 2014b). 1304.1843.

Bernstein, G. M. & Jarvis, M. *Shapes and Shears, Stars and Smears: Optimal Measurements for Weak Lensing*. *AJ*, 123:583–618 (February 2002). astro-ph/0107431.

Bertin, E. & Arnouts, S. *SExtractor: Software for source extraction. A&AS*, 117:393–404 (June 1996).

Bertone, G. & Hooper, D. *History of dark matter*. *Reviews of Modern Physics*, 90(4):045002 (October 2018). 1605.04909.

Bhattacharya, S., Habib, S., Heitmann, K., & Vikhlinin, A. *Dark Matter Halo Profiles of Massive Clusters: Theory versus Observations*. *ApJ*, 766:32 (March 2013). 1112.5479.

Blas, D., Lesgourgues, J., & Tram, T. *The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes*. *J. Cosmology Astropart. Phys.*, 7:034 (July 2011). 1104.2933.

Bleem, L. E., et al. *Galaxy Clusters Discovered via the Sunyaev-Zel'dovich Effect in the 2500-Square-Degree SPT-SZ Survey*. *ApJS*, 216:27 (February 2015). 1409.0850.

Bocquet, S., et al. *Cluster Cosmology Constraints from the 2500 deg^2 SPT-SZ Survey: Inclusion of Weak Gravitational Lensing Data from Magellan and the Hubble Space Telescope*. *ApJ*, 878(1):55 (June 2019). 1812.01679.

Bonnett, C., et al. *Redshift distributions of galaxies in the Dark Energy Survey Science Verification shear catalogue and implications for weak lensing*. *Phys. Rev. D*, 94(4):042005 (August 2016). 1507.05909.

Brammer, G. B., van Dokkum, P. G., & Coppi, P. *EAZY: A Fast, Public Photometric Redshift Code*. *ApJ*, 686(2):1503–1513 (October 2008). 0807.1533.

Bridle, S., et al. *Handbook for the GREAT08 Challenge: An image analysis competition for cosmological lensing*. *Annals of Applied Statistics*, 3:6–37 (January 2009). 0802.1214.

Buchs, R., et al. *Phenotypic redshifts with self-organizing maps: A novel method to characterize redshift distributions of source galaxies for weak lensing*. *MNRAS*, 489(1):820–841 (October 2019). 1901.05005.

Budzynski, J. M., Koposov, S. E., McCarthy, I. G., McGee, S. L., & Belokurov, V. *The radial distribution of galaxies in groups and clusters*. *MNRAS*, 423:104–121 (June 2012). 1201.5491.

Butcher, H. & Oemler, A., Jr. *The evolution of galaxies in clusters. I - ISIT photometry of C1 0024+1654 and 3C 295*. *ApJ*, 219:18–30 (January 1978).

Carlstrom, J. E., et al. *The 10 Meter South Pole Telescope*. *PASP*, 123:568–581 (May 2011). 0907.4445.

Chang, C., et al. *Modeling the Transfer Function for the Dark Energy Survey*. *ApJ*, 801:73 (March 2015). 1411.0032.

Chang, C., et al. *Dark Energy Survey Year 1 results: curved-sky weak lensing mass map*. *MNRAS*, 475:3165–3190 (April 2018a). 1708.01535.

Chang, C., et al. *The Splashback Feature around DES Galaxy Clusters: Galaxy Density and Weak Lensing Profiles*. *ApJ*, 864:83 (September 2018b). 1710.06808.

Coe, D., Benítez, N., Sánchez, S. F., Jee, M., Bouwens, R., & Ford, H. *Galaxies in the Hubble Ultra Deep Field. I. Detection, Multiband Photometry, Photometric Redshifts, and Morphology*. *AJ*, 132:926–959 (August 2006). arXiv:astro-ph/0605262.

Conroy, C., Wechsler, R. H., & Kravtsov, A. V. *Modeling Luminosity-dependent Galaxy Clustering through Cosmic Time*. *ApJ*, 647:201–214 (August 2006). astro-ph/0512234.

Cooper, M. C., et al. *The DEEP3 Galaxy Redshift Survey: Keck/DEIMOS Spectroscopy in the GOODS-N Field*. *ApJS*, 193:14 (March 2011). 1101.4018.

Costanzi, M., et al. *Methods for cluster cosmology and application to the SDSS in preparation for DES Year 1 release*. *MNRAS*, 488(4):4779–4800 (October 2019). 1810.09456.

Crocce, M., Pueblas, S., & Scoccimarro, R. *Transients from initial conditions in cosmological simulations*. *MNRAS*, 373:369–381 (November 2006). astro-ph/0606505.

Dark Energy Survey Collaboration. *Cosmology from cosmic shear with Dark Energy Survey Science Verification data*. *Phys. Rev. D*, 94(2):022001 (July 2016).

Davis, C., et al. *Cross-correlation redshift calibration without spectroscopic calibration samples in DES Science Verification Data*. *MNRAS*, 477(2):2196–2208 (June 2018). 1707.08256.

DeRose, J., et al. *The Buzzard Flock: Dark Energy Survey Synthetic Sky Catalogs*. *arXiv e-prints*, arXiv:1901.02401 (January 2019). 1901.02401.

DES Collaboration. *The Dark Energy Survey Year 3 Results: Photometric Catalog for Cosmology* (in prep a).

DES Collaboration. *The Dark Energy Survey Year 3 Results: Photometric Catalog from Deep Field Observations* (in prep b).

DES Collaboration. *The Dark Energy Survey Year 3 Results: Photometric Redshift Catalog for Cosmology* (in prep c).

DES Collaboration, et al. *The Dark Energy Survey: more than dark energy - an overview*. *MNRAS*, 460(2):1270–1299 (August 2016). 1601.00329.

DES Collaboration, et al. *Dark Energy Survey Year 1 Results: Cosmological Constraints from Cluster Abundances and Weak Lensing*. *arXiv e-prints*, arXiv:2002.11124 (February 2020). 2002.11124.

Diemer, B. & Kravtsov, A. V. *Dependence of the Outer Density Profiles of Halos on Their Mass Accretion Rate*. *ApJ*, 789:1 (July 2014). 1401.1216.

Diemer, B. & Kravtsov, A. V. *A Universal Model for Halo Concentrations*. *ApJ*, 799:108 (January 2015). 1407.4730.

Dietrich, J. P., et al. *Orientation bias of optically selected galaxy clusters and its impact on stacked weak-lensing analyses*. *MNRAS*, 443:1713–1722 (September 2014). 1405.2923.

Dietrich, J. P., et al. *Sunyaev-Zel'dovich effect and X-ray scaling relations from weak lensing mass calibration of 32 South Pole Telescope selected galaxy clusters*. *MNRAS*, 483(3):2871–2906 (March 2019). 1711.05344.

Dodelson, S., Heitmann, K., Hirata, C., Honscheid, K., Roodman, A., Seljak, U., Slosar, A., & Trodden, M. *Cosmic Visions Dark Energy: Science*. *ArXiv e-prints* (April 2016). 1604.07626.

Dodelson, S. & Schneider, M. D. *The effect of covariance estimator error on cosmological parameter constraints*. *Phys. Rev. D*, 88(6):063537 (September 2013). 1304.2593.

Drlica-Wagner, A., et al. *Dark Energy Survey Year 1 Results: The Photometric Data Set for Cosmology*. *ApJS*, 235(2):33 (April 2018). 1708.01531.

Ebeling, H., Edge, A. C., & Henry, J. P. *MACS: A Quest for the Most Massive Galaxy Clusters in the Universe*. *ApJ*, 553(2):668–676 (June 2001). astro-ph/0009101.

Eckert, K., et al. *Noise from Undetected Sources in Dark Energy Survey Images*. *arXiv e-prints*, arXiv:2004.05618 (April 2020). 2004.05618.

Efron, B. *The Jackknife, the Bootstrap and other resampling plans*. CBMS-NSF Regional Conference Series in Applied Mathematics (SIAM) (1982).

Elvin-Poole, J., et al. *Dark Energy Survey year 1 results: Galaxy clustering for combined probes*. *Phys. Rev. D*, 98(4):042006 (August 2018). 1708.01536.

Euclid Collaboration, et al. *Euclid preparation. IV. Impact of undetected galaxies on weak-lensing shear measurements*. *A&A*, 627:A59 (July 2019). 1902.00044.

Evrard, A. E., Arnault, P., Huterer, D., & Farahi, A. *A model for multiproperty galaxy cluster statistics*. *MNRAS*, 441:3562–3569 (July 2014). 1403.1456.

Farahi, A., Evrard, A. E., Rozo, E., Rykoff, E. S., & Wechsler, R. H. *Galaxy cluster mass estimation from stacked spectroscopic analysis*. *MNRAS*, 460:3900–3912 (August 2016). 1601.05773.

Farahi, A., et al. *Mass variance from archival X-ray properties of Dark Energy Survey Year-1 galaxy clusters*. *MNRAS*, 490(3):3341–3354 (December 2019). 1903.08042.

Fenech Conti, I., Herbonnet, R., Hoekstra, H., Merten, J., Miller, L., & Viola, M. *Calibration of weak-lensing shear in the Kilo-Degree Survey*. *MNRAS*, 467(2):1627–1651 (May 2017). 1606.05337.

Flaugher, B., et al. *The Dark Energy Camera*. *AJ*, 150:150 (November 2015). 1504.02900.

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. *emcee: The MCMC Hammer*. *PASP*, 125:306–312 (March 2013). 1202.3665.

Friedrich, O., Seitz, S., Eifler, T. F., & Gruen, D. *Performance of internal covariance estimators for cosmic shear correlation functions*. *MNRAS*, 456:2662–2680 (March 2016). 1508.00895.

Friedrich, O., et al. *Density split statistics: Joint model of counts and lensing in cells*. *Phys. Rev. D*, 98(2):023508 (July 2018). 1710.05162.

Gatti, M., et al. *Dark Energy Survey Year 1 results: cross-correlation redshifts - methods and systematics characterization*. *MNRAS*, 477(2):1664–1682 (June 2018). 1709.00992.

Gavazzi, G., Fumagalli, M., Cucciati, O., & Boselli, A. *A snapshot on galaxy evolution occurring in the Great Wall: the role of Nurture at z = 0*. *A&A*, 517:A73 (July 2010). 1003.3795.

Geach, J. E. & Peacock, J. A. *Cluster richness-mass calibration with cosmic microwave background lensing*. *Nature Astronomy*, 1:795–799 (November 2017). 1707.09369.

Górski, K. M., Hivon, E., Banday, A. J., Wand elt, B. D., Hansen, F. K., Reinecke, M., & Bartelmann, M. *HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere*. *ApJ*, 622(2):759–771 (April 2005). astro-ph/0409513.

Gruen, D. & Brimioulle, F. *Selection biases in empirical p(z) methods for weak lensing*. *MNRAS*, 468:769–782 (June 2017). 1610.01160.

Gruen, D., Seitz, S., Becker, M. R., Friedrich, O., & Mana, A. *Cosmic variance of the galaxy cluster weak lensing signal*. *MNRAS*, 449:4264–4276 (June 2015). 1501.01632.

Gruen, D., et al. *Weak lensing analysis of SZ-selected clusters of galaxies from the SPT and Planck surveys*. *MNRAS*, 442:1507–1544 (August 2014). 1310.6744.

Gruen, D., et al. *Density split statistics: Cosmological constraints from counts and lensing in cells in DES Y1 and SDSS data*. *Phys. Rev. D*, 98(2):023507 (July 2018). 1710.05045.

Gruen, D., et al. *Dark Energy Survey Year 1 results: the effect of intracluster light on photometric redshifts for weak gravitational lensing*. *MNRAS*, 488(3):4389–4399 (September 2019). 1809.04599.

Hansen, S. M., Sheldon, E. S., Wechsler, R. H., & Koester, B. P. *The Galaxy Content of SDSS Clusters and Groups*. *ApJ*, 699:1333–1353 (July 2009). 0710.3780.

Hao, J., Kubo, J. M., Feldmann, R., Annis, J., Johnston, D. E., Lin, H., & McKay, T. A. *Intrinsic Alignment of Cluster Galaxies: The Redshift Evolution*. *ApJ*, 740:39 (October 2011). 1103.3500.

Hastie, T., Tibshirani, R., & Friedman, J. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., New York, NY, USA (2001).

Hayashi, E. & White, S. D. M. *Understanding the halo-mass and galaxy-mass cross-correlation functions*. *MNRAS*, 388:2–14 (July 2008). 0709.3933.

Heymans, C., et al. *The Shear Testing Programme - I. Weak lensing analysis of simulated ground-based observations*. *MNRAS*, 368(3):1323–1339 (May 2006). astro-ph/0506112.

Hildebrandt, H., et al. *KiDS-450: cosmological parameter constraints from tomographic weak gravitational lensing*. *MNRAS*, 465:1454–1498 (February 2017). 1606.05338.

Hirata, C. & Seljak, U. *Shear calibration biases in weak-lensing surveys*. *MNRAS*, 343:459–480 (August 2003).

Hoekstra, H. *How well can we determine cluster mass profiles from weak lensing?* *MNRAS*, 339:1155–1162 (March 2003).

Hoekstra, H., Herbonnet, R., Muzzin, A., Babul, A., Mahdavi, A., Viola, M., & Cacciato, M. *The Canadian Cluster Comparison Project: detailed study of systematics and updated weak lensing masses*. *MNRAS*, 449:685–714 (May 2015). 1502.01883.

Hoekstra, H., Viola, M., & Herbonnet, R. *A study of the sensitivity of shape measurements to the input parameters of weak-lensing image simulations*. *MNRAS*, 468(3):3295–3311 (July 2017). 1609.03281.

Hogg, D. W. *Distance measures in cosmology*. *arXiv e-prints*, astro-ph/9905116 (May 1999). astro-ph/9905116.

Hoyle, B., et al. *Dark Energy Survey Year 1 Results: redshift distributions of the weak-lensing source galaxies*. *MNRAS*, 478(1):592–610 (July 2018). 1708.01532.

Huang, H.-J., Mandelbaum, R., Freeman, P. E., Chen, Y.-C., Rozo, E., & Rykoff, E. *Intrinsic alignment in redMaPPer clusters - II. Radial alignment of satellites towards cluster centres*. *MNRAS*, 474:4772–4794 (March 2018). 1704.06273.

Hubble, E. *A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae*. *Proceedings of the National Academy of Science*, 15(3):168–173 (March 1929).

Huff, E. & Mandelbaum, R. *Metacalibration: Direct Self-Calibration of Biases in Shear Measurement*. *ArXiv e-prints* (February 2017). 1702.02600.

Huterer, D. & Shafer, D. L. *Dark energy two decades after: observables, probes, consistency tests*. *Reports on Progress in Physics*, 81(1):016901 (January 2018). 1709.01091.

Huterer, D., et al. *Growth of cosmic structure: Probing dark energy beyond expansion*. *Astroparticle Physics*, 63:23–41 (March 2015). 1309.5385.

Jarvis, M. J., et al. *The VISTA Deep Extragalactic Observations (VIDEO) survey*. *MNRAS*, 428(2):1281–1295 (January 2013). 1206.4263.

Jarvis, M., et al. *The DES Science Verification weak lensing shear catalogues*. *MNRAS*, 460:2245–2281 (August 2016). 1507.05603.

Johnston, D. E., Sheldon, E. S., Wechsler, R. H., Rozo, E., Koester, B. P., Frieman, J. A., McKay, T. A., Evrard, A. E., Becker, M. R., & Annis, J. *Cross-correlation Weak Lensing of SDSS galaxy Clusters II: Cluster Density Profiles and the Mass–Richness Relation*. *ArXiv e-prints* (September 2007). 0709.1159.

Kacprzak, T., Bridle, S., Rowe, B., Voigt, L., Zuntz, J., Hirsch, M., & MacCrann, N. *Sérsic galaxy models in weak lensing shape measurement: model bias, noise bias and their interaction*. *MNRAS*, 441:2528–2538 (July 2014).

Kaiser, N. & Squires, G. *Mapping the Dark Matter with Weak Gravitational Lensing*. *ApJ*, 404:441 (February 1993).

Kaiser, N., Squires, G., & Broadhurst, T. *A Method for Weak Lensing Observations*. *ApJ*, 449:460 (August 1995). astro-ph/9411005.

Kaiser, N., Wilson, G., & Luppino, G. A. *Large-Scale Cosmic Shear Measurements*. *arXiv e-prints*, astro-ph/0003338 (March 2000). astro-ph/0003338.

Kannawadi, A., et al. *Towards emulating cosmic shear data: revisiting the calibration of the shear measurements for the Kilo-Degree Survey*. *A&A*, 624:A92 (April 2019). 1812.03983.

Kelly, P. L., et al. *Weighing the Giants - II. Improved calibration of photometry from stellar colours and accurate photometric redshifts*. *MNRAS*, 439:28–47 (March 2014). 1208.0602.

Kluge, M., Neureiter, B., Riffeser, A., Bender, R., Goessl, C., Hopp, U., Schmidt, M., Ries, C., & Brosch, N. *Structure of Brightest Cluster Galaxies and Intracluster Light*. *ApJS*, 247(2):43 (April 2020). 1908.08544.

Knobel, C. *An Introduction into the Theory of Cosmological Structure Formation*. *arXiv e-prints*, arXiv:1208.5931 (August 2012). 1208.5931.

Koester, B. P., McKay, T. A., Annis, J., Wechsler, R. H., Evrard, A. E., Rozo, E., Bleem, L., Sheldon, E. S., & Johnston, D. *MaxBCG: A Red-Sequence Galaxy Cluster Finder*. *ApJ*, 660:221–238 (May 2007). astro-ph/0701268.

Kormendy, J. & Bender, R. *A Proposed Revision of the Hubble Sequence for Elliptical Galaxies*. *ApJ*, 464:L119 (June 1996).

Kravtsov, A. V. & Borgani, S. *Formation of Galaxy Clusters*. *ARA&A*, 50:353–409 (September 2012). 1205.5556.

Landy, S. D. & Szalay, A. S. *Bias and variance of angular correlation functions*. *ApJ*, 412:64–71 (July 1993).

Leauthaud, A., et al. *Lensing is low: cosmology, galaxy formation or new physics?* *MNRAS*, 467:3024–3047 (May 2017).

Leauthaud, A. et al. *Lensing Without Borders I. A Blind Comparison of the Amplitude of $\Delta\Sigma$ Across Lensing Surveys* (in prep).

Lehmann, B. V., Mao, Y.-Y., Becker, M. R., Skillman, S. W., & Wechsler, R. H. *The Concentration Dependence of the Galaxy-Halo Connection: Modeling Assembly Bias with Abundance Matching*. *ApJ*, 834:37 (January 2017). 1510.05651.

Lesgourgues, J. *The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview*. *ArXiv e-prints* (April 2011). 1104.2932.

LSST Science Collaboration, et al. *LSST Science Book, Version 2.0*. *arXiv e-prints*, arXiv:0912.0201 (December 2009). 0912.0201.

MacCrann, N., et al. *DES Y1 Results: validating cosmological parameter estimation using simulated Dark Energy Surveys*. *MNRAS*, 480:4614–4635 (November 2018). 1803.09795.

MacKay, D. J. C. *Information Theory, Inference & Learning Algorithms*. Cambridge University Press, USA (2002). ISBN 0521642981.

Mandelbaum, R., et al. *GREAT3 results - I. Systematic errors in shear estimation and the impact of real galaxy morphology*. *MNRAS*, 450(3):2963–3007 (July 2015). 1412.1825.

Mandelbaum, R., et al. *The first-year shear catalog of the Subaru Hyper Suprime-Cam Subaru Strategic Program Survey*. *PASJ*, 70:S25 (January 2018). 1705.06745.

Mantz, z. B., Allen, S. W., Morris, R. G., von der Linden, A., Applegate, D. E., Kelly, P. L., Burke, D. L., Donovan, D., & Ebeling, H. *Weighing the giants- V. Galaxy cluster scaling relations*. *MNRAS*, 463:3582–3603 (December 2016). 1606.03407.

Mantz, A., Allen, S. W., Rapetti, D., & Ebeling, H. *The observed growth of massive galaxy clusters - I. Statistical methods and cosmological constraints*. *MNRAS*, 406:1759–1772 (August 2010). 0909.3098.

Mantz, A. B., et al. *Weighing the giants - IV. Cosmology and neutrino mass*. *MNRAS*, 446:2205–2225 (January 2015). 1407.4516.

Massey, R., et al. *The Shear Testing Programme 2: Factors affecting high-precision weak-lensing analyses*. *MNRAS*, 376:13–38 (March 2007). arXiv:astro-ph/0608643.

McClintock , T., Rozo, E., Becker, M. R., DeRose, J., Mao, Y.-Y., McLaughlin, S., Tinker, J. L., Wechsler, R. H., & Zhai, Z. *The Aemulus Project. II. Emulating the Halo Mass Function*. *ApJ*, 872(1):53 (February 2019). 1804.05866.

McClintock, T., Varga, T. N., et al. *Dark Energy Survey Year 1 results: weak lensing mass calibration of redMaPPer galaxy clusters*. *MNRAS*, 482(1):1352–1378 (January 2019). 1805.00039.

Medezinski, E., Broadhurst, T., Umetsu, K., Oguri, M., Rephaeli, Y., & Benítez, N. *Detailed cluster mass and light profiles of A1703, A370 and RXJ1347-11 from deep Subaru imaging*. *MNRAS*, 405:257–273 (June 2010). 0906.4791.

Medezinski, E., et al. *Planck Sunyaev-Zel'dovich cluster mass calibration using Hyper Suprime-Cam weak lensing*. *Publ. Aston. Soc. Jpn.*, 70:S28 (January 2018a). 1706.00434.

Medezinski, E., et al. *Planck Sunyaev-Zeldovich cluster mass calibration using Hyper Suprime-Cam weak lensing*. *PASJ*, 70:S28 (January 2018b). 1706.00434.

Medezinski, E., et al. *Source selection for cluster weak lensing measurements in the Hyper Suprime-Cam survey*. *Publ. Aston. Soc. Jpn.*, 70:30 (March 2018c). 1706.00427.

Medezinski, E., et al. *Source selection for cluster weak lensing measurements in the Hyper Suprime-Cam survey*. *PASJ*, 70:30 (March 2018d). 1706.00427.

Melchior, P., et al. *Mass and galaxy distributions of four massive galaxy clusters from Dark Energy Survey Science Verification data*. *MNRAS*, 449:2219–2238 (May 2015). 1405.4285.

Melchior, P., et al. *Weak-lensing mass calibration of redMaPPer galaxy clusters in Dark Energy Survey Science Verification data*. *MNRAS*, 469:4899–4920 (August 2017). 1610.06890.

Mellier, Y. *Probing the Universe with Weak Lensing*. *ARA&A*, 37:127–189 (January 1999). astro-ph/9812172.

Miller, L., Kitching, T. D., Heymans, C., Heavens, A. F., & van Waerbeke, L. *Bayesian galaxy shape measurement for weak lensing surveys - I. Methodology and a fast-fitting algorithm*. *MNRAS*, 382:315–324 (November 2007). 0708.2340.

Miller, L. et al. *Bayesian galaxy shape measurement for weak lensing surveys - III. Application to the Canada-France-Hawaii Telescope Lensing Survey*. *MNRAS*, 429:2858–2880 (March 2013). 1210.8201.

Miyatake, H., et al. *Weak-Lensing Mass Calibration of ACTPol Sunyaev-Zel'dovich Clusters with the Hyper Suprime-Cam Survey*. *ArXiv e-prints* (April 2018). 1804.05873.

Mohr, J. J., et al. *The Dark Energy Survey data management system*. In *Proc. SPIE*, volume 7016 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 70160L (July 2008). 0807.2515.

Molino, A., et al. *The ALHAMBRA Survey: Bayesian photometric redshifts with 23 bands for 3 deg*. *MNRAS*, 441(4):2891–2922 (July 2014). 1306.4968.

Mukhanov, V. *Physical Foundations of Cosmology*. Cambridge University Press (2005).

Murata, R., Nishimichi, T., Takada, M., Miyatake, H., Shirasaki, M., More, S., Takahashi, R., & Osato, K. *Constraints on the mass-richness relation from the abundance and weak lensing of SDSS clusters*. *ArXiv e-prints* (July 2017). 1707.01907.

N., M. et al. *The Dark Energy Survey Year 3 Results: Weak Lensing Shape Catalog for Cosmology* (in prep).

Navarro, J. F., Frenk, C. S., & White, S. D. M. *The Structure of Cold Dark Matter Halos*. *ApJ*, 462:563 (May 1996). astro-ph/9508025.

Noh, Y. & Cohn, J. D. *Disentangling correlated scatter in cluster mass measurements*. *MNRAS*, 426:1829–1844 (November 2012). 1204.1577.

Oaxaca Wright, C. & Brainerd, T. G. *Gravitational Lensing by NFW Halos*. *arXiv e-prints*, astro-ph/9908213 (August 1999). astro-ph/9908213.

Okabe, N. & Smith, G. P. *LoCuSS: weak-lensing mass calibration of galaxy clusters*. *MNRAS*, 461:3794–3821 (October 2016). 1507.04493.

Osato, K., Nishimichi, T., Oguri, M., Takada, M., & Okumura, T. *Strong orientation dependence of surface mass density profiles of dark haloes at large scales*. *MNRAS*, 477(2):2141–2153 (June 2018). 1712.00094.

Parzen, E. *On Estimation of a Probability Density Function and Mode*. *Ann. Math. Statist.*, 33(3):1065–1076 (09 1962).

Pereira, M. E. S., et al. *$\mu_\star$ Masses: Weak Lensing Calibration of the Dark Energy Survey Year 1 redMaPPer Clusters using Stellar Masses*. *arXiv e-prints*, arXiv:2006.10162 (June 2020). 2006.10162.

Planck Collaboration, et al. *Planck 2015 results. XXIV. Cosmology from Sunyaev-Zeldovich cluster counts*. *A&A*, 594:A24 (September 2016). 1502.01597.

Planck Collaboration, et al. *Planck 2018 results. VI. Cosmological parameters*. *arXiv e-prints*, arXiv:1807.06209 (July 2018). 1807.06209.

Postman, M. & Lauer, T. R. *Brightest Cluster Galaxies as Standard Candles*. *ApJ*, 440:28 (February 1995).

Prat, J., et al. *Dark Energy Survey year 1 results: Galaxy-galaxy lensing*. *Phys. Rev. D*, 98(4):042005 (August 2018). 1708.01537.

Press, W. H. & Schechter, P. *Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation*. *ApJ*, 187:425–438 (February 1974).

Pujol, A., Kilbinger, M., Sureau, F., & Bobin, J. *A highly precise shear bias estimator independent of the measured shape noise*. *A&A*, 621:A2 (January 2019). 1806.10537.

Reddick, R. M., Wechsler, R. H., Tinker, J. L., & Behroozi, P. S. *The Connection between Galaxies and Dark Matter Structures in the Local Universe*. *ApJ*, 771:30 (July 2013). 1207.2160.

Refregier, A. & Amara, A. *A way forward for Cosmic Shear: Monte-Carlo Control Loops*. *Physics of the Dark Universe*, 3:1–3 (April 2014). 1303.4739.

Refregier, A., Kacprzak, T., Amara, A., Bridle, S., & Rowe, B. *Noise bias in weak lensing shape measurements*. *MNRAS*, 425:1951–1957 (September 2012). 1203.5050.

Riess, A. G., et al. *A 2.4% Determination of the Local Value of the Hubble Constant*. *ApJ*, 826(1):56 (July 2016). 1604.01424.

Rowe, B. T. P., et al. *GALSIM: The modular galaxy image simulation toolkit*. *Astronomy and Computing*, 10:121–150 (April 2015). 1407.7676.

Rozo, E. & Rykoff, E. S. *redMaPPer II: X-Ray and SZ Performance Benchmarks for the SDSS Catalog*. *ApJ*, 783:80 (March 2014). 1303.3373.

Rozo, E., et al. *Cosmological Constraints from the Sloan Digital Sky Survey maxBCG Cluster Catalog*. *ApJ*, 708:645–660 (January 2010). 0902.3702.

Rozo, E., et al. *redMaGiC: selecting luminous red galaxies from the DES Science Verification data*. *MNRAS*, 461:1431–1450 (September 2016). 1507.05460.

Rykoff, E. S., Koester, B. P., Rozo, E., Annis, J., Evrard, A. E., Hansen, S. M., Hao, J., Johnston, D. E., McKay, T. A., & Wechsler, R. H. *Robust Optical Richness Estimation with Reduced Scatter*. *ApJ*, 746:178 (February 2012). 1104.2089.

Rykoff, E. S., et al. *redMaPPer. I. Algorithm and SDSS DR8 Catalog*. *ApJ*, 785:104 (April 2014). 1303.3562.

Rykoff, E. S., et al. *The RedMaPPer Galaxy Cluster Catalog From DES Science Verification Data*. *ApJS*, 224:1 (May 2016). 1601.00621.

Samuroff, S., et al. *Dark Energy Survey Year 1 results: the impact of galaxy neighbours on weak lensing cosmology with IM3SHAPE*. *MNRAS*, 475(4):4524–4543 (April 2018). 1708.01534.

Saro, A., et al. *Constraints on the richness-mass relation and the optical-SZE positional offset distribution for SZE-selected clusters*. *MNRAS*, 454:2305–2319 (December 2015). 1506.07814.

Schneider, P., van Waerbeke, L., Jain, B., & Kruse, G. *A new measure for cosmic shear*. *MNRAS*, 296:873–892 (June 1998). astro-ph/9708143.

Schrabback, T., et al. *Cluster mass calibration at high redshift: HST weak lensing analysis of 13 distant galaxy clusters from the South Pole Telescope Sunyaev-Zel'dovich Survey*. *MNRAS*, 474:2635–2678 (February 2018a). 1611.03866.

Schrabback, T., et al. *Precise weak lensing constraints from deep high-resolution K_s images: VLT/HAWK-I analysis of the super-massive galaxy cluster RCS2 J 232727.7-020437 at z = 0.70*. *A&A*, 610:A85 (March 2018b). 1711.00475.

Scoville, N. *The COSMOS 2-Degree ACS Survey*. In *HST Proposal*, page 9822 (July 2003).

Sheldon, E. *NGMIX: Gaussian mixture models for 2D images*. Astrophysics Source Code Library (August 2015). 1508.008.

Sheldon, E. S. & Huff, E. M. *Practical Weak-lensing Shear Measurement with Metacalibration*. *ApJ*, 841:24 (May 2017). 1702.02601.

Sheldon, E. S., et al. *The Galaxy-Mass Correlation Function Measured from Weak Lensing in the Sloan Digital Sky Survey*. *AJ*, 127:2544–2564 (May 2004). astro-ph/0312036.

Sifón, C., Hoekstra, H., Cacciato, M., Viola, M., Köhlinger, F., van der Burg, R. F. J., Sand, D. J., & Graham, M. L. *Constraints on the alignment of galaxies in galaxy clusters from ˜14 000 spectroscopic members*. *A&A*, 575:A48 (March 2015). 1406.5196.

Simet, M., Battaglia, N., Mandelbaum, R., & Seljak, U. *Weak lensing calibration of mass bias in the REFLEX+BCS X-ray galaxy cluster catalogue*. *MNRAS*, 466(3):3663–3673 (April 2017a). 1502.01024.

Simet, M., McClintock, T., Mandelbaum, R., Rozo, E., Rykoff, E., Sheldon, E., & Wechsler, R. H. *Weak lensing measurement of the mass-richness relation of SDSS redMaPPer clusters*. *MNRAS*, 466:3103–3118 (April 2017b). 1603.06953.

Singh, S., Mandelbaum, R., Seljak, U., Slosar, A., & Vazquez Gonzalez, J. *Galaxy-galaxy lensing estimators and their covariance properties*. *MNRAS*, 471(4):3827–3844 (November 2017). 1611.00752.

Sinha, M. & Garrison, L. *Corrfunc: Blazing fast correlation functions on the CPU*. Astrophysics Source Code Library (March 2017). 1703.003.

Smith, R. E., Peacock, J. A., Jenkins, A., White, S. D. M., Frenk, C. S., Pearce, F. R., Thomas, P. A., Efstathiou, G., & Couchman, H. M. P. *Stable clustering, the halo model and non-linear cosmological power spectra*. *MNRAS*, 341:1311–1332 (June 2003). astro-ph/0207664.

Springel, V. *The cosmological simulation code GADGET-2*. *MNRAS*, 364:1105–1134 (December 2005). astro-ph/0505010.

Springel, V., Frenk, C. S., & White, S. D. M. *The large-scale structure of the Universe*. *Nature*, 440(7088):1137–1144 (April 2006). astro-ph/0604561.

Stern, C., et al. *Weak Lensing Analysis of SPT selected Galaxy Clusters using Dark Energy Survey Science Verification Data*. *ArXiv e-prints* (February 2018). 1802.04533.

Suchyta, E., et al. *No galaxy left behind: accurate measurements with the faintest objects in the Dark Energy Survey*. *MNRAS*, 457:786–808 (March 2016). 1507.08336.

Sunyaev, R. A. & Zeldovich, Y. B. *The Spectrum of Primordial Radiation, its Distortions and their Significance*. *Comments on Astrophysics and Space Physics*, 2:66 (March 1970).

Sunyaev, R. A. & Zeldovich, Y. B. *The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies*. *Comments on Astrophysics and Space Physics*, 4:173 (November 1972).

Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. *Revising the Halofit Model for the Nonlinear Matter Power Spectrum*. *ApJ*, 761:152 (December 2012). 1208.2701.

Tanaka, M., Coupon, J., Hsieh, B.-C., Mineo, S., Nishizawa, A. J., Speagle, J., Furusawa, H., Miyazaki, S., & Murayama, H. *Photometric redshifts for Hyper Suprime-Cam Subaru Strategic Program Data Release 1*. *PASJ*, 70:S9 (January 2018). 1704.05988.

Tinker, J., Kravtsov, A. V., Klypin, A., Abazajian, K., Warren, M., Yepes, G., Gottlöber, S., & Holz, D. E. *Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality*. *ApJ*, 688:709–728 (December 2008). 0803.2706.

Tinker, J. L., Leauthaud, A., Bundy, K., George, M. R., Behroozi, P., Massey, R., Rhodes, J., & Wechsler, R. H. *Evolution of the Stellar-to-dark Matter Relation: Separating Star-forming and Passive Galaxies from z = 1 to 0*. *ApJ*, 778(2):93 (December 2013). 1308.2974.

To, C.-H., Reddick, R. M., Rozo, E., Rykoff, E., & Wechsler, R. H. *RedMaPPer: Evolution and Mass Dependence of the Conditional Luminosity Functions of Red Galaxies in Galaxy Clusters*. *arXiv e-prints*, arXiv:1910.01656 (October 2019). 1910.01656.

Troxel, M. A., et al. *Dark Energy Survey Year 1 Results: Cosmological Constraints from Cosmic Shear*. *ArXiv e-prints* (August 2017). 1708.01538.

Truemper, J. *ROSAT - A new look at the X-ray sky*. *Science*, 260:1769–1771 (June 1993).

Umetsu, K., Broadhurst, T., Zitrin, A., Medezinski, E., Coe, D., & Postman, M. *A Precise Cluster Mass Profile Averaged from the Highest-quality Lensing Data*. *ApJ*, 738:41 (September 2011). 1105.0444.

van Uitert, E., Gilbank, D. G., Hoekstra, H., Semboloni, E., Gladders, M. D., & Yee, H. K. C. *Weak-lensing-inferred scaling relations of galaxy clusters in the RCS2: mass-richness, mass-concentration, mass-bias, and more*. *A&A*, 586:A43 (February 2016). 1506.03817.

Varga, T. N., et al. *Dark Energy Survey Year 1 results: validation of weak lensing cluster member contamination estimates from P(z) decomposition*. *MNRAS*, 489(2):2511–2524 (October 2019). 1812.05116.

Varga, T. N. et al. *Dark Energy Survey Year 3 results: Synthetic Galaxy Cluster Observations* (in prep).

Voigt, L. M. & Bridle, S. L. *Limitations of model-fitting methods for lensing shear estimation*. *MNRAS*, 404(1):458–467 (May 2010). 0905.4801.

von der Linden, A., et al. *Robust weak-lensing mass calibration of Planck galaxy clusters*. *MNRAS*, 443:1973–1978 (September 2014a). 1402.2670.

von der Linden, A., et al. *Weighing the Giants - I. Weak-lensing masses for 51 massive galaxy clusters: project overview, data analysis methods and cluster images*. *MNRAS*, 439:2–27 (March 2014b). 1208.0597.

von der Linden, A. et al. *RedMapper Centering Properties. in prep.* (2018).

Wechsler, R., DeRose, J., & Busha. *ADDGALS: Simulated Sky Catalogs for Large Galaxy Surveys. in prep.* (2018).

Weinberg, S. *Cosmology* (2008).

Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., Hirata, C., Riess, A. G., & Rozo, E. *Observational probes of cosmic acceleration. Phys. Rep.*, 530:87–255 (September 2013). 1201.2434.

White, M. et al. *The Clustering of Massive Galaxies at z ˜ 0.5 from the First Semester of BOSS Data. ApJ*, 728:126 (February 2011). 1010.4915.

Yang, X., Mo, H. J., van den Bosch, F. C., Jing, Y. P., Weinmann, S. M., & Meneghetti, M. *Weak Lensing by Galaxies in Groups and Clusters - I. Theoretical Expectations. MNRAS*, 373:1159–1172 (September 2006). 0607552v2.

Zhang, Y., et al. *Dark Energy Survey Year 1 Results: Detection of Intracluster Light at Redshift 0.25. ApJ*, 874(2):165 (April 2019a). 1812.04004.

Zhang, Y., et al. *Dark Energy Surveyed Year 1 results: calibration of cluster mis-centring in the redMaPPer catalogues. MNRAS*, 487(2):2578–2593 (August 2019b). 1901.07119.

Zu, Y., Weinberg, D. H., Rozo, E., Sheldon, E. S., Tinker, J. L., & Becker, M. R. *Cosmological constraints from the large-scale weak lensing of SDSS MaxBCG clusters. MNRAS*, 439:1628–1647 (April 2014). 1207.3794.

Zuntz, J., Kacprzak, T., Voigt, L., Hirsch, M., Rowe, B., & Bridle, S. *IM3SHAPE: a maximum likelihood galaxy shear measurement code for cosmic gravitational lensing. MNRAS*, 434:1604–1618 (September 2013). 1302.0183.

Zuntz, J., et al. *Dark Energy Survey Year 1 results: weak lensing shape catalogues. MNRAS*, 481(1):1149–1182 (November 2018). 1708.01533.

Acknowledgements

I would like to thank Ralf Bender for hosting me in his research group since the time of my Master's thesis and throughout my PhD studies. There must be a parallel reality somewhere (or is it ours?), of which the University Observatory is the heaven of... I would also like to thank Stella Seitz and Daniel Gruen for their mentoring, support, useful discussions, and for keeping my spirit while navigating the perilous world of an international scientific collaboration.

Throughout the my time in the PhD I received a great amount of insight and joy from discussions with my friends and colleagues for which I am immensely grateful. Especially, I would like to mention discussions with Oliver Friedrich, Ben Hoyle and Ramona Augustin which helped me brighten my day in times when research seemed to hit a dead end. Finally I would like to thank my parents, my family and my friends for supporting me throughout my studies.